UNIVERSITÀ DEGLI STUDI DI PADOVA

DEPARTMENT OF INFORMATION ENGINEERING

Ph.D. Course in Information Engineering:
Information and Communications Technology
XXXIV Cycle

NONLINEAR PROPAGATION AND EFFECTS IN FEW-MODE FIBERS

Ph.D. Student

GIANLUCA MARCON

Supervisor

PROFESSOR MARCO SANTAGIUSTINA

January 12, 2022

NONLINEAR PROPAGATION AND EFFECTS IN FEW-MODE FIBERS

GIANLUCA MARCON

Ph.D. Course in Information Engineering:
Information and Communications Technology

January 12, 2022

Gianluca Marcon: *Nonlinear Propagation and Effects in Few-mode Fibers* ,
Ph.D. Course in Information Engineering:
Information and Communications Technology, © January 12, 2022

At times, our own light goes out and is rekindled by a spark from another person. Each of us has cause to think with deep gratitude of those who have lighted the flame within us.

— Albert Schweitzer

Dedicated to my love, Elena.

ABSTRACT

Since their birth in the mid-twentieth century, optical fibers' role in the global technological and economical landscape has become exponentially larger and more diverse, going from becoming the backbone of today's internet infrastructure, to enabling high-resolution medical imaging, or permitting the discovery of novel light sources. In each of these applications, nonlinear optical effects play an important role: in certain cases they can be the spark that lights the fire of innovation, while in others they can impose significant limitations.

In this thesis, the main results of my three-year period as a Ph.D student are presented. The first part of my research has focused on the nonlinear effect known as Raman scattering in the context of next-generation space-division multiplexed transmissions. Firstly, machine learning techniques have been applied to spectrally and spatially optimize the design of Raman amplifiers. Secondly, through the development of a theoretical model and the use of numerical simulations, the effect of fiber imperfections, such as birefringence and core ellipticity, on the gain of Raman amplifiers has been studied. The second and last part of the thesis is dedicated to the topic of Supercontinuum Generation: some early results on the characterization of novel microstructured optical fibers, achieved in collaboration with the Institute of Applied Physics (IAP) of the University of Bern, are presented.

PUBLICATIONS

Some results and figures have appeared previously in the following publications:

[1] Gianluca Marcon, Andrea Galtarossa, Luca Palmieri, and Marco Santagiustina. "Model-aware deep learning method for Raman amplification in few-mode fibers." In: *Journal of Lightwave Technology* 39.5 (2020), pp. 1371–1380.

[2] Gianluca Marcon, Andrea Galtarossa, Luca Palmieri, and Marco Santagiustina. "C+L band gain design in few-mode fibers using Raman amplification and machine learning." In: *2020 Italian Conference on Optics and Photonics (ICOP)*. IEEE. 2020, pp. 1–4.

[3] Gianluca Marcon, Andrea Galtarossa, Luca Palmieri, and Marco Santagiustina. "Gain design of few-mode fiber Raman amplifiers using an autoencoder-based machine learning approach." In: *2020 European Conference on Optical Communications (ECOC)*. IEEE. 2020, pp. 1–4.

In truth, O judges, while I wish to be adorned with every virtue, yet there is nothing which I can esteem more highly than being and appearing grateful. For this one virtue is not only the greatest, but is also the parent of all the other virtues.

– Marcus Tullius Cicero

ACKNOWLEDGMENTS

First of all, I would like to mention and thank Professor Marco Santagiustina, who has been an excellent supervisor, for giving me the opportunity to embark on this journey and guiding me during the most difficult times.

I would like to extend my gratitude also to Professor Luca Palmieri and Professor Andrea Galtarossa, who believed in me granting the financial and professional support to pursue this goal, and to Professor Alexander Heidt, who welcomed me in his laboratories of the University of Bern and gave me the opportunity to discover a new reality and meet fantastic people.

A special mention goes to the friends I made during these years. Thank you Leonardo, Riccardo, Simone, Gianluca, Marco, Luca, Francesco, Giulia, Arman, Daniele, Martina.

Then I would like to thank my parents Giancarlo and Mariella, and my sister Annachiara, for motivating me to keep my head high and never give up during these difficult years in which the world has gone upside down.

Finally, thank you Elena, for the patience you kept, for the love you showed, and for being the most important person in my life.

CONTENTS

1 INTRODUCTION 1
- 1.1 Space-division Multiplexing 1
- 1.2 Supercontinuum generation 4
- 1.3 Thesis structure 7

I RAMAN AMPLIFICATION IN FEW-MODE FIBERS FOR SPACE-DIVISION MULTIPLEXED TRANSMISSIONS

2 MACHINE LEARNING FOR RAMAN AMPLIFICATION IN FEW-MODE FIBERS 9
- 2.1 Introduction 9
- 2.2 Effects of Dataset Design 11
  - 2.2.1 Raman amplification in few-mode fibers 11
  - 2.2.2 Gain design using machine learning 12
  - 2.2.3 Dataset generation 13
  - 2.2.4 Results 14
- 2.3 Autoencoder-based Approach 18
  - 2.3.1 Proposed Method 20
  - 2.3.2 Results and validation 25
  - 2.3.3 Flat gain profiles 28
  - 2.3.4 Tilted gain profiles 31
- 2.4 Noise characteristics and sensitivity to power variation 34
  - 2.4.1 Propagation model 35
  - 2.4.2 Results 36
- 2.5 Simultaneous Gain and OSNR Optimization 40
  - 2.5.1 Proposed method 41
  - 2.5.2 Results 42
- 2.6 Conclusions 45

3 EFFECTS OF STRESS BIREFRINGENCE AND CORE ELLIPTICITY IN RAMAN-AMPLIFIED FEW-MODE FIBER LINKS 48
- 3.1 Introduction 48
  - 3.1.1 Modeling Nonlinear effects 49
- 3.2 Derivation of the model 51
  - 3.2.1 Inclusion of linear mode coupling 53
  - 3.2.2 Nonlinear terms 56
  - 3.2.3 Nonlinear polarization parameters 58
  - 3.2.4 Single-mode case 61
- 3.3 Numerical methods 63
  - 3.3.1 Linear propagation 64
  - 3.3.2 Nonlinear propagation 67

3.4 Results 68
3.4.1 Linear coupling matrices 69
3.4.2 Optimization of the nonlinear step 72
3.4.3 Simulation results 74
3.5 Conclusions 82

II SUPERCONTINUUM GENERATION

4 PRINCIPLES OF SUPERCONTINUUM GENERATION 85
4.1 Introduction 85
4.2 Fundamental processes for SC generation 87
4.2.1 Self-phase modulation 89
4.2.2 Soliton generation and fission 90
4.2.3 Influence of Raman scattering 91
4.2.4 Cross-phase modulation, four-wave mixing, and trapping of dispersive waves 92
4.3 All-normal dispersion fibers 94
4.3.1 Spectral broadening dynamics in all-normal dispersion (ANDi) fibers 96

5 CHARACTERIZATION OF SUSPENDED-CORE FIBERS 98
5.1 Fiber dispersion and multimodal characteristics 99
5.2 SC generation using a Ti:Sapphire laser 101
5.2.1 Coupling light in the suspended-core fiber (SCF) 103
5.2.2 Evolution of the supercontinuum 104
5.2.3 Polarization extinction ratio 106
5.2.4 RIN measurements 107
5.3 SC generation using an Erbium fiber laser 110
5.3.1 Polarization extinction ratio 111
5.3.2 RIN measurements 112
5.4 Conclusions 113

6 CONCLUSIONS 115
6.1 Future works 117

BIBLIOGRAPHY 118

LIST OF FIGURES

Figure 1.1 Network of undersea fiber-optic cables installed around the globe [7]. 2

Figure 1.2 Comparison of the cross-section of standard single-mode fibers (SMFs) (left) with few-mode fibers (FMFs) (center) and multi-core fibers (MCFs) (right). Adapted from [16]. 4

Figure 1.3 Scanning electron microscope images of different photonic-crystal fiber (PCF) structures [25]. 5

Figure 1.4 Comparison of the spectrum of various broadband sources [25]. 6

Figure 2.1 Diagram for the creation of Dataset B: overlap $\Delta\lambda$ between wavelength intervals for sampling the pump wavelengths is highlighted, as well the exponential distribution for the maximum allowed power for each pump, scaled by the overlap integrals of each mode. The resulting gain curves are then accepted in the dataset following the rejection sampling rules described in the text. 16

Figure 2.2 RMSE (a), MDG (b), and gain flatness (c) for the flat profiles generated using the pump parameters predicted by the neural network (NN) as a function of the target gain level. 17

Figure 2.3 Gain profiles generated using the pump parameters predicted by the NN for (a) flat and (b) tilted input target gains for different gain levels. For the tilted gain case, the total target tilt is set to 2 dB. 18

Figure 2.4 Mean (a) and variance (b) of the output of the NN during the first training iteration, as a function of the number of hidden layers and for different values of neurons per layer. 23

Figure 2.5 Diagram of the autoencoder (AE) architecture for the design of the Raman gain profile in FMFs. Green arrows and boxes are related to the training phase of the AE. 25

Figure 2.6 Attenuation profile of the considered fiber. 27

Figure 2.7 root-mean-square error (RMSE) as a function of the target gain level, for different number of pumps. Solid lines represent the mean RMSE over the 4 modes, whereas shaded areas indicate the total RMSE variation over the modes. 28
Figure 2.8 Gain flatness variation along the modes of FMF1, as a function of the target gain level, using different number of pumps. 29
Figure 2.9 Target and predicted flat gain profiles for a 4-mode fiber over the C+L–band, using 8 pumps. 30
Figure 2.10 Relative mode-dependent gain (MDG) as a function of the gain level, for different number of pumps. 30
Figure 2.11 Total pump power at $z = L$ in each mode of FMF1, as a function of the target gain level. Shaded areas indicate the variation using different number of pumps. 31
Figure 2.12 Relative MDG for flat target gain profiles, as a function of the gain level, for three different fibers. The number of pumps is set to $N_p = 8$. 32
Figure 2.13 Target and predicted gain profiles for the tilted case, using 8 pumps and with a total tilt of 1.425 dB, i.e. the maximum considered tilt during training. 33
Figure 2.14 Calculated metrics for the tilted gain case, varying the number of Raman pumps: RMSE (a)–(d), flatness (e)–(h), and MDG (i)–(l) as a function of the target gain level and target tilt. For RMSE and flatness their maximum value among the modes is reported. Columns 1 through 4 refer to the case of 4, 5, 6, and 8 pumps, respectively. 34
Figure 2.15 (a) Predicted gain for the flat case. (b) Predictions for the tilted case. (c) Close-up view of the 11 dB gain profile. 37
Figure 2.16 Mean error (a), flatness (b) and MDG (c) for flat target gains, varying the signal input power. 38
Figure 2.17 Predicted gain for the flat case when the input signal power is (a) -5 dBm and (b) -10 dBm 39
Figure 2.18 Maximum effective noise figure among modes as a function of the gain, for every signal wavelength. 39

Figure 2.19 The mode-dependent effective noise figure as a function of the Raman gain, for various signal power levels. 40
Figure 2.20 Diagram of the machine learning (ML) model for simultaneous gain and OSNR optimization. The computation of the cost function to minimize is also displayed. 41
Figure 2.21 Maximum OSNR variation and gain RMSE for the trained model, as a function of the parameter γ, for the fully counterpumping case (a) and bidirectional pumping (b). 43
Figure 2.22 Maximum OSNR variation (top) and gain flatness (bottom) over the considered spectrum, for different values of average gain levels, using different pumping schemes and values of the parameter γ. 44
Figure 2.23 Predicted flat gain profiles for the bi-directional pumping scheme with $\gamma = 0.5$ (left), and comparison of the OSNR profiles (right) between the worst-case and best-case scenarios, for an average target gain of 10 dB. 46
Figure 3.1 Nonlinear polarization parameters reported in [118] 59
Figure 3.2 Measured Raman susceptibility curves and their interpolation. Data from [119]. 61
Figure 3.3 Comparison between analytical expression and numerical solution. 63
Figure 3.4 Linear coupling matrices for the 2-modes fiber for (a) stress birefringence and (b) core ellipticity. Values are reported in logarithmic scale. 71
Figure 3.5 Linear coupling matrices for the 4-modes fiber for (a) stress birefringence and (b) core ellipticity. Values are reported in logarithmic scale. 71
Figure 3.6 Equivalent birefringence δn and core ellipticity γ for a given coupling beat length L_κ, for the two fibers considered. 72
Figure 3.7 Statistics of the overlap integrals for the two considered fibers. Represented is the maximum with respect to the index i, at the signal frequency ω_s, corresponding to $\lambda_s = 1550$ nm. 73

Figure 3.8 Mean gain of the spatial modes of the considered 2-mode fiber as a function of the coupling beat length, for two values of correlation length. Represented are the cases when signal and pump polarization are parallel (indicated with $\parallel$), and when the two are orthogonal (indicated with $\perp$), at the start of the fiber. 76

Figure 3.9 Signal power variation of the gain of the 2-mode amplifier as a function of the coupling beat length, for $L_c = 100\,\text{m}$, for parallel (left column, $\parallel$) and orthogonal (right column, $\perp$) relative polarization orientation between pump and signal. Top row: power variation of the individual modes. Bottom row: power variation on the mode groups. 77

Figure 3.10 Mean gain of the spatial modes of the considered 4-mode fiber as a function of the coupling beat length, for $L_c = 100\,\text{m}$. Represented are the cases when signal and pump polarization are parallel (indicated with $\parallel$), and when the two are orthogonal (indicated with $\perp$). 79

Figure 3.11 Signal power variation of the 4-mode amplifier as a function of the coupling beat length, for $L_c = 100\,\text{m}$, for parallel (left column, $\parallel$) and orthogonal (right column, $\perp$) polarizations between pump and signal. Top row: power variation of the individual modes. Bottom row: power variation on the "supergroups". 81

Figure 4.1 Effect of self-phase modulation (SPM) on the spectrum of a Gaussian pulse for $\phi_{max} = 40$ [26]. 89

Figure 4.2 Temporal (left) and (spectral) evolution of the fission process of a high-order soliton($N = 3$) caused by third-order dispersion. [26]. 91

Figure 4.3 Fission of a third-order soliton and corresponding red-shift due to intrapulse Raman scattering. [26]. 92

Figure 4.4 Spectrogram of the output supercontinuum generated by a 50 fs, 10 kW peak power pulse at 835 nm launched in a fiber with a zero dispersion wavelength (ZDW) of 780 nm. Adapted from [26]. 93

Figure 4.5 Typical PCF dispersion profiles for the conventional and ANDi supercontinuum (SC) generation regimes. The PCFs are preferably pumped on the anomalous dispersion region near their ZDW in conventional SC, or in the vicinity of the minimum dispersion wavelength (MDW) for ANDi SC [31]. 95
Figure 4.6 Projected axes spectrogram of (a) conventional SC and (b) ANDi SC after the broadening dynamics are concluded [31]. 95
Figure 4.7 Dynamics of SC generation in ANDi fibers. Adapted from [31]. 96
Figure 5.1 Scanning-electron microscope images of the fiber under test, showing the structure of the cladding with silica bridges supporting the core (left) and closeup of the suspended core (right). 99
Figure 5.2 Dispersion profiles of the considered SCF for input polarizations aligned ($\parallel$) or orthogonal ($\perp$) to the long core axis. 99
Figure 5.3 Mode intensity distributions and field orientations calculated using finite-element method (FEM) simulations at $\lambda = 900$. 100
Figure 5.4 Dispersion profile of the first 6 modes calculated using FEM simulations. 100
Figure 5.5 Group velocity curve of the first 6 modes calculated using FEM simulations. 101
Figure 5.6 Setup for measuring the SC and output modes. Description is given in the text. 102
Figure 5.7 Fiber input facet after core collapsing and cleaving. Photo taken from the fusion splicer display. 103
Figure 5.8 Evolution of the output spectrum as a function of the input power, when the fiber is pumped with a Ti:Sapphire laser at 810 nm with input polarization orthogonal to the long core axis. 105
Figure 5.9 Modes measured at the output of the fiber when the input polarization is orthogonal to the long-core axis of the fiber when pumped with a Ti:Sapphire laser at 810 nm, with input power 500 mW. 106

Figure 5.10 Measured polarization extinction ratio when the input polarization is parallel (left) and orthogonal (right) to the long-core axis of the fiber when pumped with a Ti:Sapphire laser at 810 nm with input power 500 mW. 106
Figure 5.11 Photos of the output SC when the fiber is pumped at 810 nm. 107
Figure 5.12 RIN spectrum (top) and corresponding integrated RIN (bottom) when the fiber is pumped using a Ti:Sapphire laser at 810 nm with input polarization orthogonal to the long core axis. The measured noise floor includes the contributions of both the photodiode and the ESA. 109
Figure 5.13 Output spectrum (top) and RIN values (bottom) measured at specific wavelengths when the fiber is pumped with a Ti:Sapphire laser at 810 nm, with input polarization orthogonal to the long core axis and input power 290 mW. 109
Figure 5.14 Evolution of the SC spectrum varying the input power, when the input polarization is parallel to the long core axis. Pump is an Erbium fiber laser centered at 1550 nm. 111
Figure 5.15 Measured polarization extinction ratio when the input polarization is orthogonal (left) and parallel (right) to the long-core axis of the fiber when pumped at 1550 nm with an Er:fiber laser at 290 mW input power. 112
Figure 5.16 Modes captured at the fiber output at 1550 nm for input pump polarization (a) orthogonal and (b) parallel to the long core axis, when pumped with an Er:fiber laser with input power = 290 mW. 112
Figure 5.17 RIN spectrum (top) and corresponding integrated RIN (bottom) when the fiber is pumped using a Er:fiber laser at 1550 nmnm. 113
Figure 5.18 Output spectrum (top) and RIN values (bottom) measured at specific wavelengths when pumped with an Erbium fiber laser centered at 1550nm, with input polarization parallel to the long core axis and input power 290 mW. 114

LIST OF TABLES

Table 2.1 Values of the overlap integrals for the considered FMF. 15

Table 2.2 Overlap integrals of the FMFs used for simulation, in units of $1 \times 10^9\,\mathrm{m}^{-2}$. 26

Table 2.3 Values of $I_{m,n}$ as defined in Eq. (2.2) (units of $10^9\,\mathrm{m}^{-2}$). 36

Table 3.1 Expression of the overlap integrals that define the strength of the nonlinear interaction between modes. 58

1 INTRODUCTION

Since the advent of the laser in 1960 [1], its application to communications technology seemed an obvious use. The thirst for bandwidth by telecommunications carriers, justified by the growing telephone and television traffic, demanded new methods to transmit information at higher rate and longer distances. Optical frequencies were hundreds of terahertz, and promised a far greater bandwidth than the microwave technologies that were used at the time. The shift to laser transmission then seemed natural, although finding suitable media over which the optical signals could be transmitted proved to be a serious challenge, as the signal attenuation introduced by the materials were prohibitively high. In fact, when Ivan Kaminow proposed the use of optical fibers as waveguides to transmit light, the best available glass at the time, silica, showed losses of about $1\,\mathrm{dB\,m^{-1}}$ [2]. Thanks to the 2009 Nobel Laureate Charles Kuen Kao and his visionary ideas, the main sources of attenuation were identified in the glass purity, rather than on fundamental physical phenomena such as scattering [3]; this meant that with the appropriate manufacturing practices silica impurities could be removed and ultra-low losses could be achieved. In 1978, a few years after glass manufacturer Corning demonstrated the possibility of producing optical fibers with losses of $7\,\mathrm{dB\,km^{-1}}$ [4], researchers at Nippon Telegraph and Telephone (NTT) were able to make fibers achieving attenuation coefficients as low as $0.2\,\mathrm{dB\,km^{-1}}$ at the wavelength of 1550 nm, close the the fundamental limit imposed by Rayleigh scattering [5].

Since then, optical fibers have been shown to be the perfect medium for long-distance optical communications, and to this day constitute the backbone of the world-wide telecommunications infrastructure. With more than 4 billion kilometers of fibers installed worldwide, and a projected market value of about $7 billion, fiber optic technology is one of the most important enablers of today's digital society [6].

1.1 SPACE-DIVISION MULTIPLEXING

During the 4 previous decades, thanks to the research efforts in discovering all-optical amplification devices in erbium-doped fiber amplifiers (EDFAs), new multiplexing techniques such as wavelength-division multiplexing (WDM), and high-spectral efficiency coherent modula-

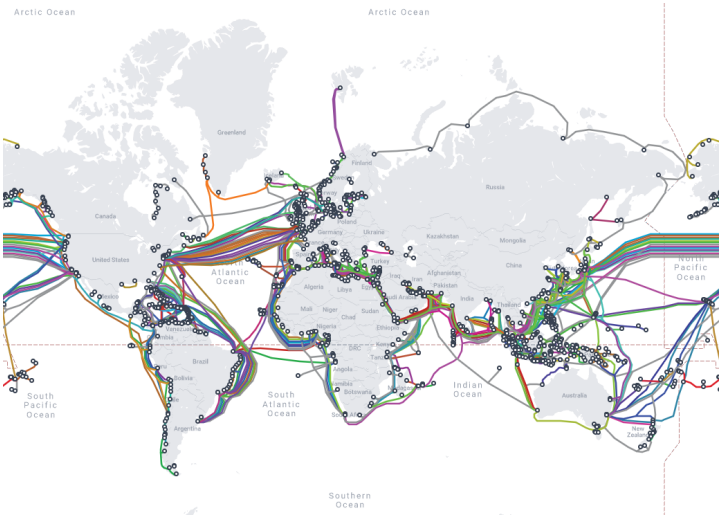


Figure 1.1: Network of undersea fiber-optic cables installed around the globe [7].

tion formats, the capacity of standard single-mode fibers (SMFs) cables has been increasing by an order of magnitude every 4 years [8].

This outstanding growth is however bound to be stopped. A fundamental limit to the transmission capacity of silica optical fibers exists, and is dictated by the material's nonlinear properties. In fact, nonlinear phenomena such as self-phase modulation (SPM), cross-phase modulation (XPM), and four-wave mixing (FWM) stemming from silica's Kerr coefficient, can accumulate over long propagation distances and introduce an additional source of noise in the system. As the effects are enhanced when the transmitted power is increased, an increase in the number of parallel channels transmitted in a SMF is met with a reduction of its total capacity, defining what is known as the nonlinear Shannon limit (NSL) of the nonlinear fiber optic channel, which sets a theoretical upper bound of SMFs' transmission capacity at about $100\,\mathrm{Tbit\,s^{-1}}$ when filling the entire C+L optical band at a spectral efficiency of $10\,\mathrm{bit\,s^{-1}\,Hz^{-1}}$. [9, 10]. Experimentally, the record capacity for SMF systems has reached $206.1\,\mathrm{Tbit\,s^{-1}}$, achieved on a 54 km span with distributed amplification over the S+L+C bands and a spectral efficiency of approximately $11\,\mathrm{bit\,s^{-1}\,Hz^{-1}}$ [11].

Over the last 20 years, internet traffic experienced a growth similar to Moore's law for transistor densities, reaching a compound annual growth rate (CAGR) of 40% over two decades, doubling every 2 years [12]. This increase in bandwidth demand is set to continue in the coming years, with Cisco forecasting an annual Internet Protocol traffic of 4.8 ZB per year by 2022, with machine-to-machine (M2M) connections accounting for more than half of the global connected devices [13]. With these enormous growth rates, bandwidth demands will soon

reach the upper limit of the current optical fiber infrastructure, causing what has been known as the "capacity crunch" [14].

Although the deployment of new SMF links could in practice be a viable solution to sustain the ever-growing demands in internet traffic, this approach would not be favorable in terms of cost and energy-efficiency scaling, which would depend linearly on the number of newly installed links. The main candidate technology for next-generation optical communications is instead found in space-division multiplexing (SDM) [10], which, as its name suggests, aims at exploiting the space of a single optical fiber cable as an additional dimension over which parallel information streams can be transmitted simultaneously.

To this end, new types of optical fibers that can offer spatially-multiplexed transmissions have been proposed as transmission medium: few-mode fibers (FMFs) and multi-core fibers (MCFs). While MCFs enable spatial diversity through the physical separation of different information streams across multiple cores embedded in a single cladding structure, FMFs maintain a single-core geometry with a larger core-radius compared to SMFs, supporting the propagation of multiple guided configurations of the electromagnetic field (modes) over which independent data streams can be transmitted (fig. 1.2). Although the concept of using multimode fibers (MMFs) for data transmission has in fact preceded that of SMFs, its use to date is limited to short reach applications such as in data-centers, where the ability to support multiple modes is not exploited to increase transmission capacity, but rather for ease of deployment.

Together with the increase in transmission capacity, these new fibers bring a new set of challenges with them. One the major issues from a technological standpoint consists in the crosstalk between the different spatial channels, which results in an additional OSNR penalty at the receiver and in an increase of the computational cost of equalization algorithms. Although it can arise due to the vicinity of the individual cores of a MCF, crosstalk in FMFs has its origins in the coupling between modes as a result of imperfections (residual birefringence, ellipticity of the core, etc.) or external perturbations (bending, twist, etc.) breaking the ideal cylindrical symmetry of the fiber geometry [15].

Although it might appear that nonlinear effects have played an exclusively antagonistic role in the framework of optical communications, they also enabled several key technologies, with fiber-optical parametric amplifiers [17], all-optical switching [18], and optical phase-conjugation [19] being only a handful of examples of what nonlinear fiber optics has made possible. Among the different types of nonlinear effects arising in optical fibers, stimulated Raman scattering (SRS)

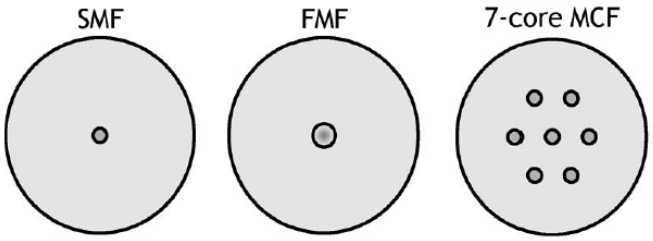


Figure 1.2: Comparison of the cross-section of standard SMFs (left) with FMFs (center) and MCFs (right). Adapted from [16].

has been particularly important for the realization of new all-optical amplification schemes [20]. In fact, single-mode Raman amplifiers have been thoroughly researched since the late 90s, and have enabled dramatic increases in the reach and capacity of optical communication systems thanks to their distributed nature enabling lower noise figures compared to conventional EDFAs, and the possibility of using polychromatic pumping schemes to achieve arbitrarily wide amplification bands [21].

In the context of SDM, a great deal of research is performed to design systems and components that can be efficiently integrated with a favorable cost, energy, and capacity scaling [10]. For SDM Raman amplifiers, and particularly for FMF systems, these requirements translate to the necessity of minimizing the mode-dependent gain and guaranteeing the flexibility and favorable characteristics demonstrated for SMF links under the new challenges brought by modal crosstalk, while leveraging the economic appeal guaranteed by component reduction in pump-sharing circuits [22].

1.2 SUPERCONTINUUM GENERATION

The field of nonlinear optics has also benefited greatly from Charles Kao's work in reducing the attenuation of optical fibers. Before then, since its birth with the invention of the laser in 1960, there was little interest in using optical fibers as a nonlinear medium, as the already small nonlinear coefficient of silica glass was further hindered by its high attenuation. Traditional nonlinear optics revolved around the use of bulk materials and crystals, which permitted the discovery of second-harmonic generation, SRS, stimulated Brillouin scattering (SBS), and FWM [23]. Bulk nonlinear optics was however limited by practical issues such as focusing, beams overlap, and self-focusing, which could easily cause damage to the material [24].

With the first low-loss fibers produced, the low nonlinearity of silica was more than made up by the great increase in the length over which

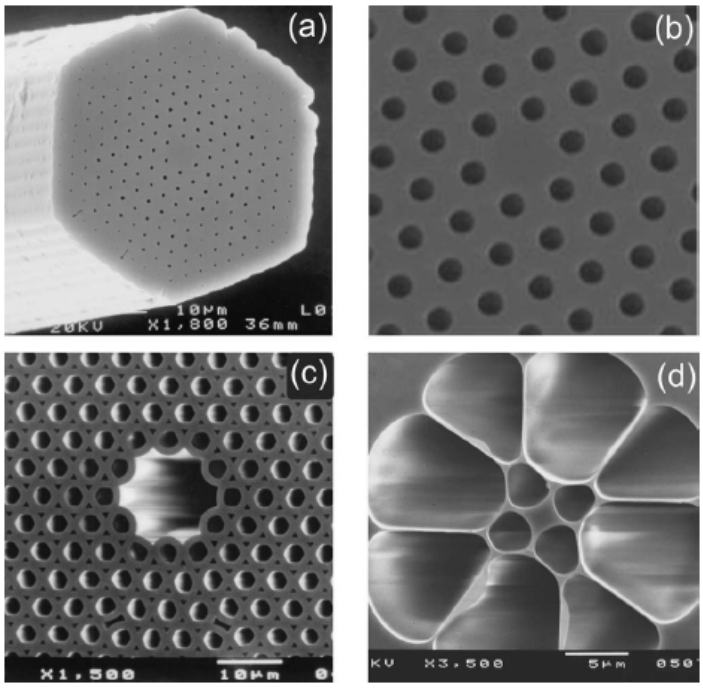


Figure 1.3: Scanning electron microscope images of different PCF structures [25].

high optical intensities could be maintained, which was also possible thanks to the small spot size of single-mode fibers. Coupled with the far higher practicality of using an optical waveguide, the newly formed field of nonlinear fiber optics began to flourish throughout the 70s and the 80s [23, 24]. With improvements in fiber manufacturing processes in the late 1990s, a new type of optical waveguide, formed by a periodic lattice of air holes embedded in a silica cladding, called photonic-crystal fiber (PCF) was demonstrated in 1996 by Philip Russell [25] (fig. 1.3).

The guiding of light through a PCF is not based on conventional index-guiding principles, but rather on a *photonic bandgap* created by its Bragg-like periodic structure. This breakthrough in fiber-optics enabled the possibility of guiding light through a core of materials with a much lower index of refraction than the surrounding cladding, potentially even air. The main benefits of using such fibers for application in nonlinear optics came from two major features, namely the possibility of confining light on a much smaller core than usual, thus increasing the optical power density, and the ability to tailor the dispersion properties of the waveguide through an accurate design of the fiber's structure in order to take full advantage of the available high-power sources [23].

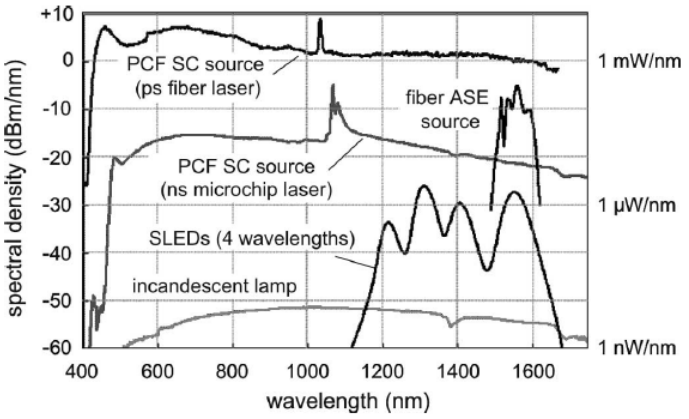


Figure 1.4: Comparison of the spectrum of various broadband sources [25].

With advances in research, several new types of microstructured optical fibers were developed, allowing for unprecedented nonlinear efficiencies on interaction lengths of just a few centimeters. The umbrella term that is aptly used for these novel waveguides is highly-nonlinear fibers (HNLFs). One of the most successful applications of PCFs, and HNLFs in general, is the so-called supercontinuum (SC) generation. When high-power, pico- or femtosecond laser pulses are injected in the fiber, the interplay of different nonlinear effects such as SPM, SRS, FWM and the dispersive properties of the fiber cause a massive broadening of the input spectrum (fig. 1.4).

A new type of light source, characterized by high brightness, high coherence, and with spectral widths of several hundreds to few thousands of nanometers was born [26]. In the years, SC sources have found widespread applications in several different fields where coherent broadband light is required, such as optical coherence tomography [27], optical frequency metrology [28], spectroscopy [29], and even in the simultaneous generation of WDM channels for optical communications [30].

In this context, research has focused on a particular type of HNLFs characterized by a an all-normal dispersion (ANDi) profile. The main benefits of these fibers consist in the generation of uniform and flat SCs, a higher noise-suppression capability, and pulse-preservation characteristics, which can be important in applications in ultra-fast photonics or time-resolved spectroscopy to generate pulses in the order of a few optical cycles with a high degree of coherence [31].

1.3 THESIS STRUCTURE

In this thesis the main results achieved during my three-year Ph.D. program will be presented. In order to give a clearer description of the different topics that have been the focus of my research, the structure of the thesis will be organized as follows:

1. Part i is dedicated to the study of Raman amplification in few-mode fibers for application in space-division multiplexed systems:
    - In chapter 2, methods based on machine learning are presented and applied to the case of distributed Raman amplification schemes in order to spectrally and spatially optimize their gain in few-mode fibers supporting up to 6 groups of linearly polarized modes.
    - In chapter 3, an analytical and numerical model that describes the dynamics of Raman amplification in few-mode fibers in the presence of linear mode coupling caused by the fiber manufacturing process is derived and presented. With this tool, the statistic on the amplifier gain and on the output signal power fluctuations are studied in different regimes of linear coupling and for fibers supporting 2 and 4 groups of linearly polarized modes.
2. Part ii is dedicated the topic of supercontinuum generation:
    - In chapter 4, an overview of the various nonlinear phenomena that permit the massive spectral broadening in optical fibers is given, discussing the main differences between different supercontinuum generation regimes.
    - In chapter 5, preliminary results of the experimental characterization of novel microstructured ANDi suspended-core fibers, achieved in collaboration with the Institute of Applied Physics (IAP) of the University of Bern (Switzerland), and the Leibniz Institute of Photonic Technology, Jena (Germany) are presented.
3. Finally, chapter 6 concludes the thesis, summarizing the main results obtained during the Ph.D program, and highlighting the possible future developments on the topics that have been treated in this document.

Part I

RAMAN AMPLIFICATION IN FEW-MODE FIBERS FOR SPACE-DIVISION MULTIPLEXED TRANSMISSIONS

2

MACHINE LEARNING FOR RAMAN AMPLIFICATION IN FEW-MODE FIBERS

2.1 INTRODUCTION

As introduced in Chapter 1, the capacity of current optical fiber infrastructure based on SMFs is rapidly approaching the NSL imposed by the Kerr nonlinearity [32], which coupled with the ever-increasing demand of internet traffic during the last few decades, will result in a "capacity crunch" [14]. SDM has emerged as the leading technology to overcome the NSL, exploiting the spatial diversity of MCFs and FMFs to transmit multiple data streams in a single optical fiber cable, thus offering a cost-effective, energy-efficient, and scalable solution [10]. In FMFs, SDM is implemented by using the orthogonality of the guided modes to transmit independent data streams, thereby realizing mode-division multiplexing (MDM) [33].

In order to benefit from the added capacity of spatially-multiplexed transmissions, suitable network devices must be designed fully compatible with the already well-established techniques such as WDM. To this end, the role of SDM-compatible amplifiers is of fundamental importance, with several experimental works demonstrating the effectiveness in MDM scenarios of EDFAs [34, 35] and Raman amplifiers (RAs) in increasing the system capacity in both long-haul and short reach scenarios [36–40].

The compensation of link losses with minimal signal-to-noise ratio (SNR) reduction has always been a crucial aspect in optical communications, but additional care must be taken with SDM systems to minimize both mode-dependent gain (MDG) and wavelength-dependent gain (WDG), as they can be both detrimental to the multiple-input multiple-output (MIMO) digital signal processing (DSP) algorithms that mitigate the effect of mode-crosstalk to correctly recover the transmitted signals [38].

While the simplicity and power efficiency of EDFAs made them appealing for commercial communication systems, their reduced gain bandwidth has made Raman amplification an attractive solution for wideband WDM schemes [21]. The spectral flexibility of RAs, together with suitable optimization techniques, enables the design of flat gain profiles over large bandwidths by means of multiple wavelength pumps [21]. In the context of SDM, the additional degrees of freedom can lead to higher control of WDG and MDG [41]. Addition-

ally, RAs can offer distributed amplification, resulting in a reduced noise contribution compared to EDFAs [42].

Different approaches have been proposed to optimize the gain of multi-pump SMF RAs [43–45], however, they generally require complex iterative algorithms involving multiple integrations of the propagation equations and must be executed for each target gain spectrum. This aspect is particularly important for FMFs, as the number of parameters to optimize increases with the number of modes of the fiber [46].

A recent publication [47] proposed a new machine learning (ML) technique to solve this problem. Specifically, a neural network (NN) can be trained to learn the inverse relationship $\mathbf{y} = f^{-1}(\mathbf{G})$ between the vector $\mathbf{y}$ of pump wavelengths and powers and the corresponding gain profile $\mathbf{G}$, using a synthetic dataset $\mathcal{D} = \{(\mathbf{y}_i, \mathbf{G}_i)\}$ of thousands of gain curves generated with random pump parameters. The learned mapping is then used to compute the required pump parameters $\tilde{\mathbf{y}} = \widetilde{f^{-1}}(\mathbf{G}_{target})$ to approximate a given target gain profile. This eliminates the need to solve complex iterative algorithms that require multiple integrations of the propagation equations for every new target profile, which instead is replaced with a single evaluation of the trained NN at a negligible computational cost making this method an attractive solution for control-planes of next-generation self-adaptive and autonomous optical networks, where low-latency, near real-time operation is needed [47].

The authors of [47] used two additional techniques to refine the prediction of the NN. The first is *model-averaging*, which consists in training several NNs in parallel, each on a random permutation of the dataset, and finally averaging their output. This approach, while providing some improvements, is heavier in terms of computational time, both for the training and the inference phase.

The second technique consists in a fine-tuning phase requiring an additional NN trained to learn the direct mapping $\mathbf{G} = f(\mathbf{y})$. The prediction error on the gain profile obtained with the approximate pump parameters $\tilde{\mathbf{y}}$ is estimated using the learned direct mapping $\tilde{f}$ and minimized using an iterative gradient-descent algorithm without integrating the propagation equations. Publication [47] showed promising results, demonstrating the feasibility of the method with flat and tilted gain profiles using a counter-propagating RA over the C and C+L bands, achieving a maximum prediction error on the considered gain profiles well below 1 dB for different levels of amplifications.

Since the publication of Ref. [47], this line of research has produced several other contributions involving the use of NNs for Raman amplification in SMFs, including experimental demonstrations of both Raman and hybrid Raman/EDFA schemes [48–50], the simultaneous

gain design and corresponding noise figure prediction [51], and the optimization of both the spectrum and the evolution of the Raman gain along the fiber [52].

In the context of MDM, a similar approach to design flat gain profiles both for 2-mode and 4-mode fibers has been demonstrated in [53]. This work does not use neither model-averaging nor fine-tuning algorithms; therefore, computational requirements for the inference phase are reduced. For the 4-mode FMF, [53] showed encouraging results in terms of MDG and gain flatness; nevertheless, the analysis considers only flat gain profiles, is limited to the C band, and does not provide accuracy metrics that take the average amplification level into account, meaning that it is not known if the transmitted signals are either over- or under-amplified.

2.2 EFFECTS OF DATASET DESIGN

While the optimization of the pump parameters in FMFs can be improved employing more sophisticated unsupervised (i.e. that do not rely on a dataset of $(\mathbf{G}, \mathbf{y})$ pairs for training) ML schemes [54, 55] to obtain high accuracy predictions on the C+L bands and using fibers with higher mode counts, as will be shown in the following sections of this chapter, in this section we show that the performance of the method in [53] can be increased if careful design choices are made during the dataset generation phase. The ability of the NN to generate tilted gain spectra, which can be important to e.g., offset the Raman scattering-induced tilt of the transmitted signals, or compensate the wavelength-dependent losses introduced by filters or multiplexers, is also tested [47].

2.2.1 *Raman amplification in few-mode fibers*

In a few-mode RA supporting M modes, N_s signal wavelengths and N_p pump wavelengths, the power evolution of the ith frequency propagating in the mth mode is described by the following set of nonlinear ordinary differential equations [46, 56]:

$$\begin{aligned}\eta_i \frac{dP_i^m}{dz} = &-\alpha_i P_i^m \\ &+ P_i^m \sum_{j=i+1}^{N_s+N_p} \sum_{n=1}^{M} \mathcal{I}_{m,n} g_R(|f_i - f_j|) P_j^n \\ &- P_i^m \sum_{j=1}^{i-1} \sum_{n=1}^{M} \frac{f_i}{f_j} \mathcal{I}_{m,n} g_R(|f_i - f_j|) P_j^n,\end{aligned} \tag{2.1}$$

where P_i^m is the power in the m-th mode and i-th frequency, where $i \in \{1, \ldots, N_s + N_p\}$, $m \in \{1, \ldots, M\}$; α_i is the attenuation coefficient at the ith frequency and $g_R(\Delta f)$ is the Raman gain coefficient for the frequency shift Δf. Frequencies f_i are sorted in increasing order, meaning that frequency f_j with $j > i$ acts as a pump for all frequencies f_k with $k < i$.

The strength of the nonlinear interaction is given by the coefficients $\mathcal{I}_{m,n}$, i.e the overlap integrals between the mth and nth mode, defined by

$$\mathcal{I}_{m,n} = \frac{\iint_{-\infty}^{+\infty} F_m(x,y)F_n(x,y)dxdy}{\iint_{-\infty}^{+\infty} F_m(x,y)dxdy \iint_{-\infty}^{+\infty} F_n(x,y)dxdy}, \tag{2.2}$$

where $F_k(x,y)$ is the intensity profile of the kth mode, which is assumed to be wavelength-independent on the considered bandwidth [56]. Finally, η_i determines the relative propagation direction of the i-th frequency, so for the counter-propagating pumps $\eta_i = -1$, $\forall i \in \{N_s + 1, \ldots, N_s + N_p\}$, whereas $\eta_i = 1$ for the first N_s frequencies. Modes with similar propagation constants, i.e. those within the same mode group, exhibit high coupling efficiency, resulting in the equalization of the amplifier gain for that particular group. For the purpose of RA they can simply be treated as a unique mode [46, 57]. Conversely, linear mode coupling between different mode groups is weak and will be neglected here, like in [46, 53, 57]. In a RA of length L, the on-off gain $\mathbf{G} = [G_i^m]$ is defined as

$$G_i^m = \frac{P_i^m(z = L) \text{ with pumps turned on}}{P_i^m(z = L) \text{ with pumps turned off}}, \tag{2.3}$$

where $i = 1, \cdots, N_s$.

2.2.2 *Gain design using machine learning*

If we consider the function $\mathbf{G} = f(\mathbf{y})$ that computes the Raman on-off gain for a given vector $\mathbf{y} = [\mathbf{P}, \boldsymbol{\lambda}]$ of pump powers and wavelengths, the optimization of the gain profile is achieved if the inverse function $\hat{\mathbf{y}} = f^{-1}(\hat{\mathbf{G}})$ is determined for the given target $\hat{\mathbf{G}}$. The method presented in [47] for SMFs and later extended in [53] for the FMF case aims at obtaining an approximation of $f^{-1}(\cdot)$ over a wide region of the space of possible Raman gain profiles using a NN. Starting from a dataset $\mathcal{D} = \{(\mathbf{G}_i, \mathbf{y}_i)\}_{i=1}^D$ of gain profiles and corresponding pump

parameters, we can train the NN by finding its parameters θ that solve the optimization problem

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \frac{1}{D} \sum_{i=1}^{D} \mathcal{L}\left(\mathbf{y}_i, \mathcal{N}(\mathbf{G}_i\,;\theta)\right), \tag{2.4}$$

where $\mathcal{N}(\cdot\,;\theta)$ is the NN and $\mathcal{L}(\cdot,\cdot)$ is a cost function. In practice, (2.4) is generally solved using iterative gradient-based algorithms such as stochastic gradient descent (SGD) or adaptive moment estimation (Adam) [58]. After training, the NN approximates the inverse function f^{-1}, and can be queried on gain profiles that are not present in the dataset to obtain the corresponding vector of pump parameters with negligible computational cost.

2.2.3 *Dataset generation*

In order to approximate the function f^{-1} on a given region of its domain (i.e. the region of approximately flat and tilted profiles), it is essential to train the NN using a dataset that accurately samples that particular region. In the case of Raman amplifiers, where the dataset is constructed by solving (2.1) with random values of the pump parameters, the space of gain profiles is sampled indirectly through the function f. The selection of the pump parameters is therefore critical, as it determines the quality of the available data. This aspect is particularly important for wide-band RAs, where the number of Raman pumps increases and the space of pump parameters gets bigger, and even more so when considering the propagation on multiple spatial channels.

In the dataset generation phase, as in [53], the pump power is uniformly sampled from a certain interval $P_{i,m} \sim \mathcal{U}(0, P_{i,m}^{\max})$. However, instead of setting the same maximum power for each wavelength, here we propose to prioritize shorter wavelength pumps (which experience more depletion) by introducing an exponential allocation scheme that divides the total maximum power per mode P_{tot} on the ith wavelength according to

$$P_{i,m}^{\max}(\alpha) = P_{\text{tot}} \cdot \frac{\exp[-\alpha(i-1)]}{\sum_{j=1}^{N_p} \exp[-\alpha(j-1)]}, \tag{2.5}$$

where we assumed that wavelengths are sorted in ascending order, and α is a design parameter.

Additionally, in order to account for the different pumping efficiency of each mode of the fiber, we can scale the maximum pump power as follows

$$\bar{P}_{i,m}^{\max} = P_{i,m}^{\max} \cdot \frac{\bar{\mathcal{I}}}{\mathcal{I}_{m,m}}, \quad \bar{I} = \max_n \mathcal{I}_{n,n}, \tag{2.6}$$

where $\mathcal{I}_{m,n}$ are the intensity overlap integrals defined in (2.2).

In [47] and [53], the pump wavelengths are uniformly sampled from disjoint intervals of a selected portion of the optical band Δ_λ. This choice is forced to impose an ordering of outputs of the NNs, which is necessary for the training algorithm effectively minimize the cost function. However, with this approach many gain profiles are excluded from the dataset, so we consider the case in which the partitions of Δ_λ partially overlap, trading a small penalty on the training accuracy for a more complete description of the space of gains.

Finally, to reduce the number of gain profiles that lie outside the selected region of interest (and that can possibly reduce the performance of the NNs) we can apply rejection sampling to discard them. A possible scheme could consist in accepting only gain profiles with an average gain in a certain interval, or with an MDG lower than a selected threshold. More stringent rules for the acceptance of the generated gains result in a higher rejection probability, and thus increase simulation time.

2.2.4 *Results*

The proposed scheme is tested on a FMF with core diameter of 14 µm, core refractive index of 1.46, and a relative refractive index difference between core and cladding equal to 0.28%, supporting the propagation of the LP01 and LP11 mode groups. Its intensity overlap integrals are reported in Table 2.1. The Raman gain spectrum of the fiber is computed using the method detailed in [59] and setting the peak gain coefficient to $7 \times 10^{-14}\,\mathrm{m\,W^{-1}}$. The transmitted signals consist in $N_s = 50$ wavelengths equally spaced on the C+L optical band, each carrying $-10\,\mathrm{dBm}$ per mode, while $N_p = 8$ co-propagating Raman pumps are used.

After a hyperparameter optimization phase consisting of a grid-search algorithm on the number of layers and neurons per layer, and considering different activation functions, we selected a feed-forward NNs using 5 fully-connected linear layers of 1000 neurons each. Batch normalization and parametric rectified linear unit (PReLU) activation functions are applied to the output of each layer, with the exception of the output layer. Batch normalization is used to normalize the output of each layer and facilitate the training of the NNs [58].

Table 2.1: Values of the overlap integrals for the considered FMF.

$\mathcal{I}_{m,n}$	LP_{01}	LP_{11}
LP_{01}	$7.01 \times 10^{-9}\,m^{-2}$	$3.95 \times 10^{-9}\,m^{-2}$
LP_{11}	$3.95 \times 10^{-9}\,m^{-2}$	$5.67 \times 10^{-9}\,m^{-2}$

The predictions of the same NN architecture trained on two different datasets are compared: "Dataset A" is generated using a total maximum pump power of 800 mW per mode uniformly divided between wavelengths, and sampling the pump wavelengths from non-overlapping partitions of the interval Δ_λ from 1410 nm to 1525 nm; "Dataset B" is instead computed with the exponential allocation scheme with $\alpha = 0.3$ and scaling the total maximum pump power of each mode according to (2.6), while wavelengths are sampled from partitions of Δ_λ overlapping by 10 nm. In both cases, we apply rejection sampling to exclude gain profiles with an average gain outside of the interval from 5 dB to 15 dB, and with a total gain variation over the C+L band greater than 5 dB. The described setting is illustrated in fig. 2.1, although with a lower number of pumps for the sake of clarity.

The datasets contain 20000 gain curves each, of which 80% is dedicated to training the NNs, while the remaining 20% is used for validation. An additional pre-processing step is applied to the generated datasets, normalizing each gain and corresponding pump parameter vector to speed up the convergence of the training algorithm.

The NNs are trained for 500 epochs, using the Adam algorithm with a learning rate $\eta = 5 \times 10^{-4}$ to update their weights and the mean squared error cost function. In order to reduce overfitting and improve performance on gain curves outside the training dataset, *dropout* regularization is employed, randomly dropping the connections between fully-connected layers with a probability $p = 0.2$ at each training iteration [60].

After training, the two neural networks are evaluated on flat gain profiles with different levels of average gain. The predicted pump parameters are used to solve (2.1) and obtain the predicted on-off gain $\tilde{\mathbf{G}}$ using (2.3). We then test the accuracy of the predictions by calculating their root-mean-square error (RMSE) with respect to the input targets for each mode m as

$$\mathrm{RMSE}_m(\mathbf{G}, \tilde{\mathbf{G}}) = \sqrt{\frac{1}{N_s} \sum_{i=1}^{N_s} \left(G_i^m - \tilde{G}_i^m\right)^2}. \tag{2.7}$$

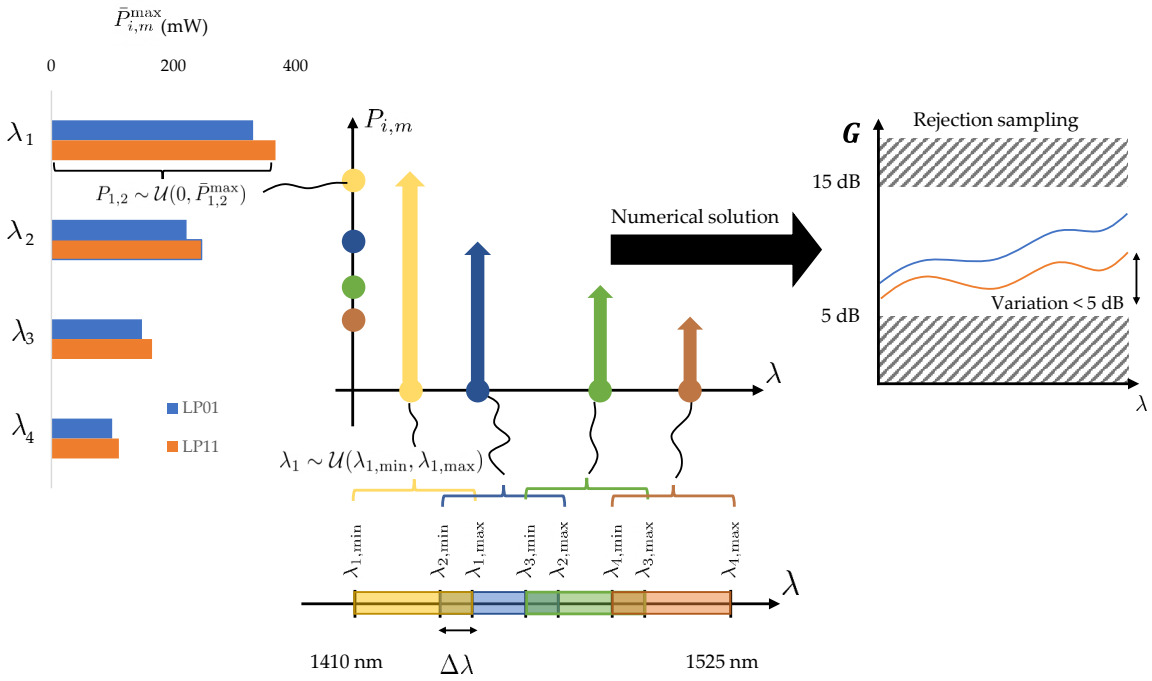


Figure 2.1: Diagram for the creation of Dataset B: overlap $\Delta\lambda$ between wavelength intervals for sampling the pump wavelengths is highlighted, as well the exponential distribution for the maximum allowed power for each pump, scaled by the overlap integrals of each mode. The resulting gain curves are then accepted in the dataset following the rejection sampling rules described in the text.

The resulting RMSE is reported for both datasets in Fig. 2.2 (a) for different levels of amplification, showing that for Dataset B the error is always smaller and stays below 0.4 dB, while the curve for Dataset A quickly increases for low and high gains.

For a given gain profile $\mathbf{G}$, its MDG and flatness can be evaluated according to the following equations

$$\mathrm{F}_m(\mathbf{G}) = \max_i G_i^m - \min_i G_i^m, \tag{2.8}$$

$$\mathrm{MDG}(\mathbf{G}) = \max_i \left(\max_m G_i^m - \min_m G_i^m \right). \tag{2.9}$$

Applying them to the predicted gain profiles, we can make similar considerations to the RMSE case both for the MDG in Fig. 2.2 (b), where for Dataset B the curve is fairly constant and always lower than 0.3 dB, and for gain flatness in Fig. 2.2 (c), with a maximum value of approximately 0.8 dB at a target gain level of 15 dB.

Finally, gain profiles obtained using the pump parameters predicted by NN trained on Dataset B, along with the corresponding target

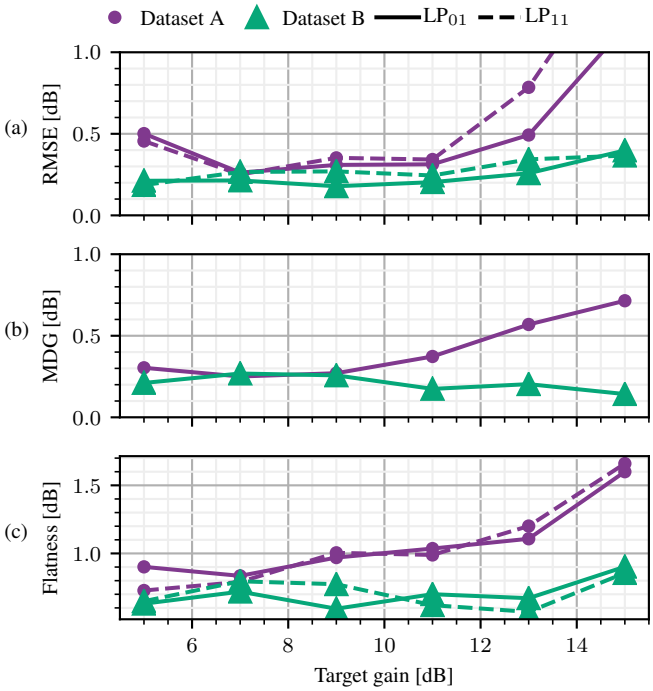


Figure 2.2: RMSE (a), MDG (b), and gain flatness (c) for the flat profiles generated using the pump parameters predicted by the NN as a function of the target gain level.

curves, are reported in Fig. 2.3 (a) for different levels of amplifications, showing the good accuracy achieved by the method. In Fig. 2.3 (b) we also show the predicted gain profiles when interrogating the NN with target gains with a total tilt of 2 dB on the considered bandwidth, with similar accuracy to the flat profile case.

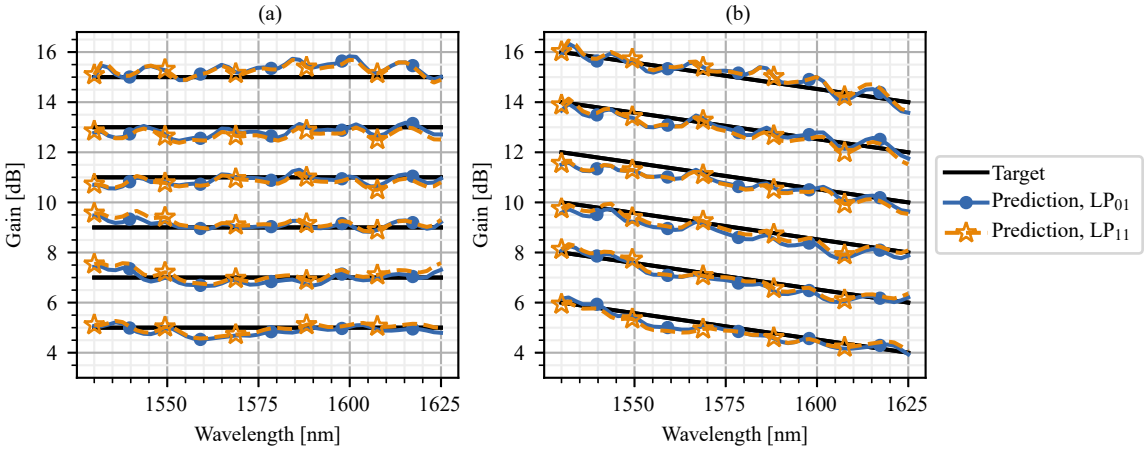


Figure 2.3: Gain profiles generated using the pump parameters predicted by the NN for (a) flat and (b) tilted input target gains for different gain levels. For the tilted gain case, the total target tilt is set to 2 dB.

2.3 AUTOENCODER-BASED APPROACH

The main drawback of the methods of Refs. [47, 53] (and by extent, the one reported in the previous section of this document) with respect to iterative optimization algorithms, is that while the latter specifically look at minimizing a cost function $\mathcal{C}(\mathbf{G}_{target}, \tilde{\mathbf{G}})$ between a desired and predicted gain profile by taking the propagation model into account, the former is instead optimized to minimize a cost function $\mathcal{C}(\mathbf{y}, \tilde{\mathbf{y}})$ between pump parameters. The NNs are thus unaware of the underlying mathematical and physical relations between pump parameters and gain profile, which has to be learned from the available data. In order to approximate the inverse function $\mathbf{y} = f^{-1}(\mathbf{G})$ using a NN and generate flat gain profiles, the region of space of approximately flat Raman gains, must be properly sampled. This cannot be easily achieved since the training dataset is generated by solving the Raman equations with randomly drawn pump powers and wavelengths, meaning that only the codomain of $f^{-1}(\cdot)$ is sampled with full control. As a result, only a minor part of the generated gains fall inside the region of interest, resulting in the NNs being trained to learn the inverse function on a much bigger domain than required, potentially hindering its performance on flat/tilted gains. This aspect is also more problematic when increasing the amplification bandwidth or the number of modes and pumps, as the dimensionality of the space to explore also increases. As seen in section 2.2, the choice of parameters for the generation of the dataset is critical for the effectiveness of these methods. In fact, the powers and wavelengths of the pumps are selected a priori, meaning

that the trained NNs might not able to predict the optimal pump parameters if the preliminary selection of the dataset parameters are not accurate.

Owing to automatic differentiation (AD) techniques [61], analytical or numerical models describing dynamical systems can be embedded in ML architectures [62]. By recording the series of elementary operations performed on the model input in a computational graph, AD libraries such as PyTorch [63] can compute the exact derivatives of the model output with respect to any parameter to be optimized [64]. In the context of optical communications, this approach has been demonstrated to be able to perform end-to-end (E2E) optimization of a intensity modulation/direct detection system by jointly optimizing the transmitter and receiver using NNs, outperforming classical feed-forward equalization [65]. The effectiveness of this technique has also been demonstrated for coherent transmissions [66–69] where probabilistic constellation shaping and geometric constellation shaping have shown to be fundamental for achieving record spectral efficiencies in short- and long-haul experiments [70].

In this section an unsupervised ML method based on autoencoders (AEs) is proposed, employing AD to embed a differentiable FMF Raman amplification model in the training procedure of a NN to predict the pump parameters able to generate flat and tilted gain profiles over a pre-determined range of amplification levels and gain tilts. The trained NN can then be used to obtain the required pump parameters for a desired gain profile with low time-complexity. The presented method has the advantage to train the NN directly on the searched (e.g. flat and tilted) gain profiles, thereby directly sampling the selected region of space of possible gains, instead of building a dataset by solving the Raman equations using random pump parameters. The supervised dataset design phase, along with the issues related to it, is thus completely avoided, with the relationship between target gain and pump parameters being learned in the training phase of the NN through the differentiable Raman model. The ability to directly target an arbitrary region of the space of Raman gains makes this method easily generalizable to any type of gain profiles, more robust and scalable with respect to the changes in number of modes, Raman pumps, and fiber parameters.

2.3.1 *Proposed Method*

2.3.1.1 *Machine Learning Model Architecture*

Many of the E2E learning methods in the literature are based on a ML architecture called AE [58]. An AE is composed of two main blocks: an encoder,

$$\mathcal{E}(\,\cdot\,;\theta_e): \quad \mathbb{R}^p \to \mathbb{R}^q, \tag{2.10}$$

and a decoder,

$$\mathcal{D}(\,\cdot\,;\theta_d): \quad \mathbb{R}^q \to \mathbb{R}^p, \tag{2.11}$$

where θ_e, θ_d are learnable parameters and $q < p$. The role of the encoder is to learn a lower dimensionality representation $\hat{\mathbf{x}}$ of its input data $\mathbf{x}$ in a way that enables the decoder to compute an estimate $\tilde{\mathbf{x}}$ of the original data from $\hat{\mathbf{x}}$:

$$\tilde{\mathbf{x}} = \mathcal{D}(\mathcal{E}(\mathbf{x};\theta_e);\theta_d). \tag{2.12}$$

Typically, both $\mathcal{E}$ and $\mathcal{D}$ consist in NNs that are jointly trained to minimize the average of the cost function $\mathcal{C}_{\text{AE}}$ between original and reconstructed samples of a dataset $\mathcal{X} = \{\mathbf{x}_i\}$:

$$\theta_e^*, \theta_d^* = \underset{\theta_e,\theta_d}{\operatorname{argmin}} \; \frac{1}{|\mathcal{X}|} \sum_{\mathbf{x}\in\mathcal{X}} \mathcal{C}_{\text{AE}}\big(\mathcal{D}(\mathcal{E}(\mathbf{x};\theta_e);\theta_d), \mathbf{x}\big). \tag{2.13}$$

By replacing the decoder $\mathcal{D}$ with a differentiable Raman model $\mathcal{R}$ that maps a vector of pump powers and wavelengths

$$\mathbf{y} = [\boldsymbol{\lambda} \mid \mathbf{P}] \in \mathbb{R}_+^{(M+1)N_p}, \tag{2.14}$$

to the corresponding on-off gain, we can train the AE using (2.13) on a dataset $\mathcal{X} = \{\mathbf{G}_i\}$ of gain curves to force the encoder NN to learn a low-dimensional representation that minimizes the reconstruction error through $\mathcal{R}$. That is, the trained encoder approximates the inverse of the Raman model

$$\mathcal{E}(\,\cdot\,;\theta_e^*) \approx \mathcal{R}^{-1}(\,\cdot\,), \tag{2.15}$$

meaning that the lower dimensionality representation of the input gain $\mathbf{G}$ is the vector $\mathbf{y}$ of pump powers and wavelengths that approximates it.

While the numerical integration of the Raman model $\mathcal{R}$ is still required in the forward-pass of the training process to compute (2.13), this computational cost is no longer needed to determine the pump parameters that approximate a target gain profile, which are directly obtained by using $\mathcal{E}(\,\cdot\,;\theta_e^*)$.

Here, the encoder $\mathcal{E}$ is a feed-forward (FF) NN with N_h hidden, fully connected (FC) layers of N_n neurons and rectified linear unit (ReLU) activation functions. Input and ouput layers have size $N_s \times M$ and $N_p \times (M+1)$, respectively, where the $M+1$ factor accounts for the fact that for each pump M pump power values and 1 wavelength value has to be predicted.

In order to force a constraint on the predicted pump parameter vector $\mathbf{y}$, a sigmoidal function

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{2.16}$$

is used to limit the output $\mathbf{x}$ of the last FC layer of the NN to the open interval $(0,1)$, as is customarily done in the ML framework [60]. The resulting normalized pump vector $\hat{\mathbf{y}}$ can then be linearly mapped to the desired interval of powers and wavelengths.

The decoder $\mathcal{R}$ consists of a fixed-step, fourth-order Runge-Kutta integrator that solves (2.1) to compute the on-off gain using the pump parameters generated by the encoder.

2.3.1.2 *Training algorithm*

The optimal encoder parameters, θ_e^*, are found by solving (2.13) with an iterative training algorithm and using the RMSE between target and approximated gain as a cost function:

$$\mathcal{C}_{\mathrm{AE}}(\mathbf{G}, \tilde{\mathbf{G}}) = \frac{1}{M} \sum_{m=1}^{M} \underset{i}{\mathrm{RMSE}} \left(G_i^m, \tilde{G}_i^m \right), \tag{2.17}$$

for $i = 1, \ldots, N_s$. In the k-th iteration of the training algorithm, the AE reconstruction of each curve in the dataset $\mathcal{X}$ is computed as

$$\tilde{\mathbf{G}} = \mathcal{E}(\mathcal{R}(\mathbf{G}); \theta_e(k)) \quad \forall \mathbf{G} \in \mathcal{X}, \tag{2.18}$$

where $\theta_e(k)$ are the encoder parameters at the current iteration. The total cost function for the iteration is then evaluated by averaging (2.17) over $\mathcal{X}$

$$\mathcal{C}(k) = \frac{1}{|\mathcal{X}|} \sum_{\mathbf{G} \in \mathcal{C}} \mathcal{C}_{\mathrm{AE}}(\mathbf{G}, \tilde{\mathbf{G}}). \tag{2.19}$$

Finally, the encoder parameters are updated with a gradient descent algorithm

$$\theta_e(k+1) = \theta_e(k) - \epsilon \nabla_{\theta_e(k)} \mathcal{C}(k), \tag{2.20}$$

where $\epsilon > 0$ is the learning rate (LR) of the algorithm. The exact computation of the gradients is performed by means of AD and backpropagation [58]. Advanced optimization algorithms such as the

Adam algorithm [71] are typically employed for the update step (2.20) as they offer robust convergence properties and adaptive LR schemes for each parameter.

During training, the relationship between input target gains and their respective pump powers and wavelengths are learned through the differentiable Raman solver $\mathcal{R}$, meaning the vector of pump parameters $\mathbf{y}_i$ associated to each gain profile $\mathbf{G}_i$ of the dataset is not needed. This fact can be exploited by completely bypassing the dataset generation phase and training the encoder on the targeted family of desired ideal gain profiles.

Here the focus is on flat and tilted gain spectra, so in each training iteration k we generate a batch $\mathcal{B}_k = \{\mathbf{G}_i\}_{i=1}^{B}$ of B ideal gain profiles with average gain level l_G and tilt t_G (gain variation per unit wavelength) randomly sampled from a uniform distribution

$$l_G \sim \mathcal{U}\left(l_G^{\min}, l_G^{\max}\right), \tag{2.21}$$

$$t_G \sim \mathcal{U}\left(t_G^{\min}, t_G^{\max}\right), \tag{2.22}$$

It is important to notice that this approach is completely generalizable and not limited to flat and tilted gains only, but it could be extended to other families of gain profiles by properly including them in the training data.

As previously detailed, in supervised learning techniques such as those presented in [47] and [53], the underlying physical model is only described by and learned from the provided data, meaning that it is essential to use datasets that are representative of the problem. In the context of RA, this means that the dataset must properly sample the region of possible Raman gains containing approximately flat gain profiles in order for the NN to properly learn the inverse Raman model. This cannot be done efficiently or easily, as there is actually no direct control on which gain profiles are sampled, but rather on the power and wavelength of each pump. Instead, the presented approach avoids this issue by directly sampling the selected space of Raman gains. Consequently, the problem of overfitting is completely avoided, and regularization techniques are not required.

2.3.1.3 *Initial conditions*

When training the AE using the algorithm described above we face the problem of local minima, which is common when dealing with the optimization of many parameters with complex cost functions. An important aspect to consider when dealing with local minima is the initial conditions of the algorithm, which can significantly affect the outcome of the optimization problem.

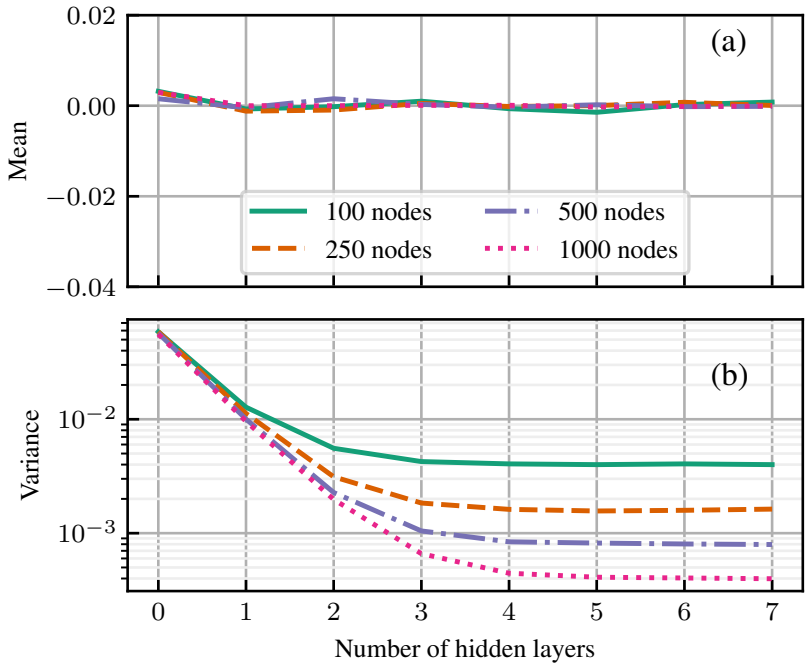


Figure 2.4: Mean (a) and variance (b) of the output of the NN during the first training iteration, as a function of the number of hidden layers and for different values of neurons per layer.

The parameters of the encoder's FC layers are initialized by sampling a uniform distribution on the interval $[-\sqrt{n}, \sqrt{n}]$, where n is the inverse of the number of incoming connections to that layer [63]. This approach has been demostrated to be effective to mitigate the problem of *vanishing gradients* when training multi-layer NNs [72]. While this random initialization strategy is beneficial in classic supervised learning models, it affects the initial condition of our AE, as it imposes a random value to the initial normalized pump parameter vector $\hat{\mathbf{y}}_\mathbf{0}$. We analyzed the statistical distribution of the output of the last FC layer, $\mathbf{x}_\mathbf{0}$, during the first training iteration for different number of hidden layers and neurons. We found that its elements follow a Gaussian-like distribution with zero mean and a variance that decreases as the number of layers and neurons increases. In Fig. 2.4 (a) and (b) we show the mean and variance, respectively, of $\mathbf{x}_\mathbf{0}$ for the case of a 4-mode fiber with 50 wavelength channels and 8 pumps, resulting into an input layer of 200 neurons and an output layer of 40 neurons. For a sufficiently high number of hidden layers and neurons (which is easily met in practice) we can then use the approximation $\mathbf{x}_\mathbf{0} \approx \mathbf{0}$, meaning that by (2.16) in the first training iteration $\hat{\mathbf{y}}_\mathbf{0} \approx \sigma(\mathbf{0}) = \mathbf{0.5}$, so fixing the initial pump powers and wavelengths to the middle point

of the interval of allowed values. By introducing a centering vector $\boldsymbol{\alpha}$ and subtracting it to the input of the sigmoids, we have:

$$\hat{\mathbf{y}} = \sigma(\mathbf{x} - \boldsymbol{\alpha}), \tag{2.23}$$

which enables us to relate the initial pump parameters to $\boldsymbol{\alpha}$ as follows

$$\hat{\mathbf{y}}_0 = \sigma(\mathbf{x}_0 - \boldsymbol{\alpha}) \approx \sigma(-\boldsymbol{\alpha}) = \frac{1}{1 + e^{\boldsymbol{\alpha}}}. \tag{2.24}$$

We can use this result to force a more desirable initial condition on the pump parameters by computing the appropriate value of $\boldsymbol{\alpha}$ by inverting (2.24).

2.3.1.4 *Counter-propagating pumps*

For the case of counter-propagating pumps it is customary to implement a differential equation solver based on a shooting algorithm to determine the correct initial pumps powers to be injected at $z = L$. This however would require significantly more computational resources, as the propagation equation should be solved several times for each training sample and, more importantly, could introduce convergence problems [73].

However, the method proposed here presents a particularly advantageous feature on this regard: in fact, the encoder $\mathcal{E}$ can direcly predict the pumps powers at $z = 0$, $\tilde{P}_i^m(z = 0)$, eliminating the need to employ shooting algorithms. By solving (2.1) with initial ($z = 0$) conditions for pumps and signals, we obtain the predicted gain $\tilde{\mathbf{G}}$ along with pumps powers at the end of the link, $\tilde{P}_i^m(z = L)$, which are the values of interest. We therefore trade a significant computation cost in the training phase for a single integration of (2.1) in the inference phase. The resulting AE-based system is represented in the diagram of Fig. 2.5, highlighting the various components of the architecture and its input-output relations. Green boxes and arrows are related to the training phase of the AE, during which the encoder parameters θ_e are optimized. In order to compensate the significantly higher sensitivity of the predicted gain to the optimization parameters and avoid further problems with local minima, we introduce a modification to the training algorithm by multiplying the output of the last FC layer $\mathbf{x}$ by a mask $\mathbf{H}_k$, where the subscript k indicates the k-th training iteration. The normalized pump parameters for the k-th training iteration are then determined by:

$$\hat{\mathbf{y}}_k = \sigma(\mathbf{x}_k \odot \mathbf{H}_k - \boldsymbol{\alpha}), \tag{2.25}$$

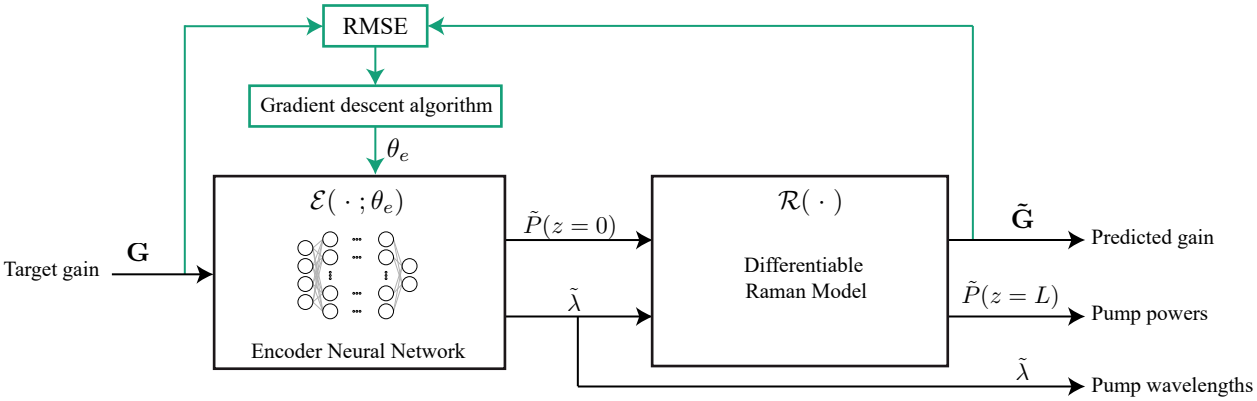


Figure 2.5: Diagram of the AE architecture for the design of the Raman gain profile in FMFs. Green arrows and boxes are related to the training phase of the AE.

where $\odot$ indicates the Hadamard (element-by-element) product. $\mathbf{H}_k$ can be suitably designed to "steer" the NN by weighting the computed gradients of the cost function with respect to the pump parameters during backpropagation. In our case, we set:

$$\mathbf{H}_k = [H_k^i] = \begin{cases} 0 & 1 \leq i \leq N_p\,,\ k < K \\ 1 & N_p + 1 \leq i \leq (M+1)N_p\,,\ k < K \\ 1 & 1 \leq i \leq (M+1)N_p\,,\ k \geq K\,, \end{cases} \tag{2.26}$$

where the superscript i indicates the i-th element of the vector. Using this definition, the pump wavelengths are fixed to their initial conditions for the first K iterations, allowing the encoder to learn just the relationship between predicted pump power and generated gain profile, which is more critical during the first training iterations. The training algorithm is summarized in Algorithm 1 and is completely implemented using the PyTorch ML library [63], which enables us to leverage AD and graphics processing unit (GPU) acceleration.

2.3.2 *Results and validation*

We test the presented method using counter-propagating pumps and a $L = 70$ km long 4-mode step-index fiber (SIF) whose overlap integrals are calculated in [46] and reported in Table 2.2. Hereinafter, we refer to this fiber as FMF1. The Raman gain spectrum is computed using the multi-vibrational-mode model of the Raman response function for silica fibers [59], whereas the peak value for the Raman gain coefficient $g_R = 7 \times 10^{-14}\,\mathrm{W^{-1}\,m}$ was used [42]. The spectral attenuation coefficient of the fiber is obtained from a parabolic fit of attenuation

Algorithm 1 AE training algorithm

Compute centering vector $\boldsymbol{\alpha}$
Initialize encoder parameters: $\theta_e(0)$
for $k = 0$ **to** $N_{\text{iter}} - 1$ **do**
  Compute the mask $\mathbf{H}_k$
  Generate batch $\mathcal{B}_k = \{\mathbf{G}_i\}_{i=1}^{B}$ of gain profiles
  Propagate batch to obtain the pump parameters from NN:
  $\hat{\mathbf{Y}}_k = \{\hat{\mathbf{y}}_k^i\} = \{\sigma(\hat{\mathbf{x}}_k^i \odot \mathbf{H}_k - \boldsymbol{\alpha})\}$, $\hat{\mathbf{x}}_k^i = NN(\mathbf{G}_i; \theta_e(k))$
  Map the normalized parameters to the selected range:
  $\mathbf{Y}_k = \{\mathbf{y}_k^i\} = \text{Scale}(\hat{\mathbf{Y}}_k)$
  Integrate (2.1) to compute the predicted gain profiles:
  $\tilde{\mathcal{B}}_k = \{\tilde{\mathbf{G}}_i^k\} = \{\mathcal{R}(\mathbf{y}_k^i)\}$
  Compute the cost function $\mathcal{C}(k)$ using (2.19)
  Compute gradients with backpropagation: $\nabla_{\theta_e(k)}\mathcal{C}(k)$
  Update the parameters $\theta_e(k+1)$ using (2.20)
end for

Table 2.2: Overlap integrals of the FMFs used for simulation, in units of $1 \times 10^9\,\text{m}^{-2}$.

	FMF1				FMF2			
	LP_{01}	LP_{11}	LP_{02}	LP_{21}	LP_{01}	LP_{11}	LP_{02}	LP_{21}
LP_{01}	6.24	4.12	4.62	2.85	5.47	3.6	3.87	2.45
LP_{11}	4.12	4.36	2.33	3.81	3.6	5.7	1.95	3.28
LP_{02}	4.62	2.33	6.15	2.12	3.87	1.95	4.94	1.76
LP_{21}	2.85	3.81	2.12	3.88	2.45	3.28	1.76	4.95

data of a commercially available SMF, $\alpha(\lambda) = \alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2$, with coefficients $\alpha_0 = 5.788\,\text{dB}\,\text{km}^{-1}$, $\alpha_1 = -7.1246 \times 10^{-3}\,\text{dB}\,\text{km}^{-1}\,\text{nm}^{-1}$, $\alpha_2 = 2.268 \times 10^{-6}\,\text{dB}\,\text{km}^{-1}\,\text{nm}^{-2}$. This approximation is valid for the entire range of considered wavelengths, and is reported in Fig. 2.6. As in [46, 56] we assume the absence of mode-dependent losses (MDL).

We consider the transmission on $N_s = 50$ wavelengths on the C+L band, for a total number of spatial and wavelength channels equal to $N_{ch} = M \times N_s = 200$. The input power for each channel is set to $P_{ch} = -10$ dBm.

The encoder NN is composed of $N_h = 5$ hidden layers of $N_n = 1000$ neurons each, and its parameters are optimized using the Adam algorithm with a LR $\epsilon = 1 \times 10^{-4}$. The AE is trained for $N_{\text{iter}} = 1000$ iterations with batches of $B = 1024$ gain curves, which are sufficient to fill the GPU random access memory (RAM) and ensure 100% GPU

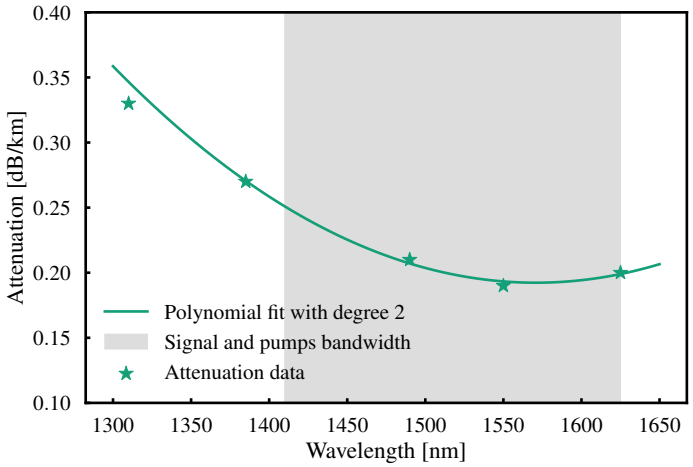


Figure 2.6: Attenuation profile of the considered fiber.

clock utilization. Each batch is generated according to the strategy described in section 2.3.1, with average gain level and tilt uniformly sampled from the intervals of 5 dB to 15 dB and $-0.015\,\mathrm{dB\,nm^{-1}}$ to $0.015\,\mathrm{dB\,nm^{-1}}$, respectively.

We map the output of the sigmoids to limit the predicted power at $z = 0$ and wavelength of each pump into the intervals $I_{P(z=0)} = [-60, 20]$ dBm and $I_\lambda = [1410, 1520]$ nm, respectively.

Using (2.24) we set the initial power on each pump to $P_0(z = 0) = 3$ dBm, whereas the wavelengths are uniformly distributed over I_λ. Additionally, we use (2.26) to fix the pumps wavelengths to their initial value for the first $K = 100$ iterations.

Once the AE is trained, the encoder is used determine pump wavelengths and powers at $z = 0$ to approximate a given target gain profile:

$$\tilde{\mathbf{y}} = [\tilde{\boldsymbol{\lambda}} \mid \tilde{\mathbf{P}}(z = 0)] = \mathcal{E}(\mathbf{G}\,;\,\theta_e^*), \tag{2.27}$$

and the corresponding predicted gain $\tilde{\mathbf{G}}$ and pumps powers at $z = L$ are obtained with a single integration of the Raman equations (2.1):

$$[\tilde{\mathbf{G}} \mid \tilde{\mathbf{P}}(z = L)] = \mathcal{R}(\tilde{\mathbf{y}}). \tag{2.28}$$

The total training time for the employed NNs is approximately 45 minutes using an NVIDIA Quadro M4000 GPU. The computational time to perform a prediction for a single target gain profile on an Intel consumer laptop CPU is approximately 11 ms, of which 1 ms is required for computing the output of the encoder NN, and the remaining 10 ms are needed for integrating (2.1).

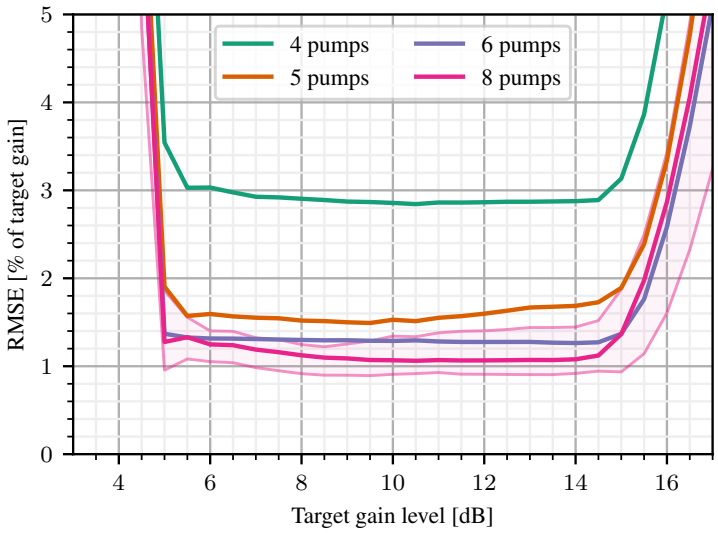


Figure 2.7: RMSE as a function of the target gain level, for different number of pumps. Solid lines represent the mean RMSE over the 4 modes, whereas shaded areas indicate the total RMSE variation over the modes.

2.3.3 *Flat gain profiles*

First, we assess the performance of the presented method using FMF1 for the case of flat target gain profiles in terms of RMSE, gain flatness and MDG, for different levels of amplification and varying the number of Raman pumps. Given that the number of pumps determines the size of the input and output layers of the encoder NN, the training algorithm must be run for every value that this parameter assumes. For each target curve, we obtain the corresponding AE prediction using (2.27), (2.28) and compute the RMSE for each mode m as in eq. (2.7). In Figure 2.7 we report the RMSE in terms of percentage of the target gain level, as a function of the amplification level and using 4, 5, 6, and 8 Raman pumps. Solid lines and shaded regions represent the average RMSE and maximum to minimum RMSE variation over the modes, respectively. For gain levels inside the target interval of $[5, 15]$ dB, the RMSE curves are almost constant, independently of the number of pumps used. Conversely, the RMSE rapidly grows outside the training interval, as the encoder NN is not able to extrapolate the correct pump parameters. By increasing the number of Raman pumps from 4 to 8 we improve the RMSE, going from 3% to about 1% of the target gain.

A clear picture on the improvements brought by an increased number of pumps is given by the gain flatness, defined for each mode as in eq. (2.9). The gain flatness in reported in Fig. 2.8 in terms of percentage of target gain level, where the shaded areas represent the

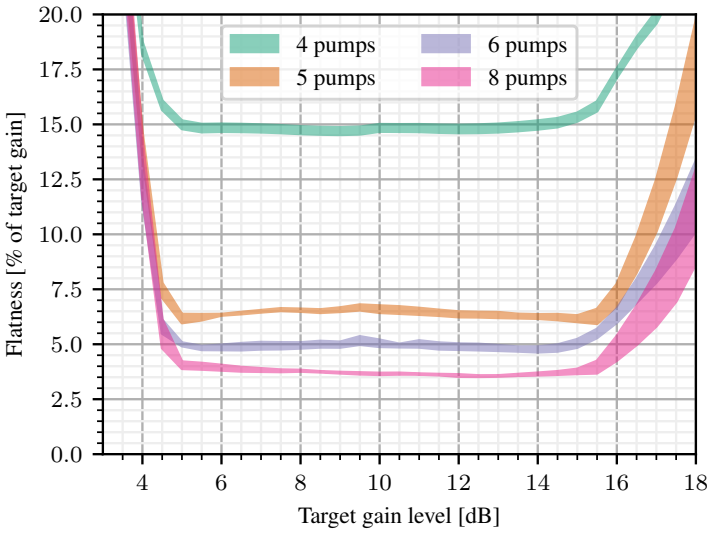


Figure 2.8: Gain flatness variation along the modes of FMF1, as a function of the target gain level, using different number of pumps.

flatness variation over the modes. The most significant improvement is obtained from 4 to 5 pumps, reducing the flatness from 15% to about 6%. For example, this means that for a 10 dB target gain, the total flatness would be decreased to just 0.6 dB from 1.5 dB; this value is further decreased to 0.35 dB using 8 pumps. Moreover, we can observe that flatness is practically constant among the modes, with fluctuations always lower than 0.5% of the target gain in the interval from 5 dB to 15 dB. We can see an example of the achieved gain profiles for the case of 8 pumps in Fig. 2.9, where we plot the flat target profiles and the predicted gain curves for different amplification levels inside the training interval. The gain profile for each mode is in fact the same up to a residual MDG, which increases with the gain level.

For a given gain profile $\mathbf{G}$, we quantify its MDG as detailed in eq. (2.8). In Fig. 2.10 we report the MDG as percentage of the total gain using 4, 5, 6, and 8 Raman pumps. Differently from the case of gain flatness, the number of pumps does not influence the total MDG, which is practically constant inside the interval of gain levels on which the AE was trained, settling at about 2% of the target gain. This residual MDG is caused mainly by the fact that LP_{01} and LP_{11} modes are systematically over-amplified with respect to the others. By inspecting the values of overlap integrals of FMF1 in Table 2.2, we can observe that the sum of the off-diagonal entries in the columns/rows associated with LP_{01} and LP_{11} are the first and second largest, respectively, meaning that power is more efficiently coupled by the nonlinear Raman interaction in these two modes. In Fig. 2.11 we plot the total pump power in $z = L$ on each mode of the FMF, as predicted

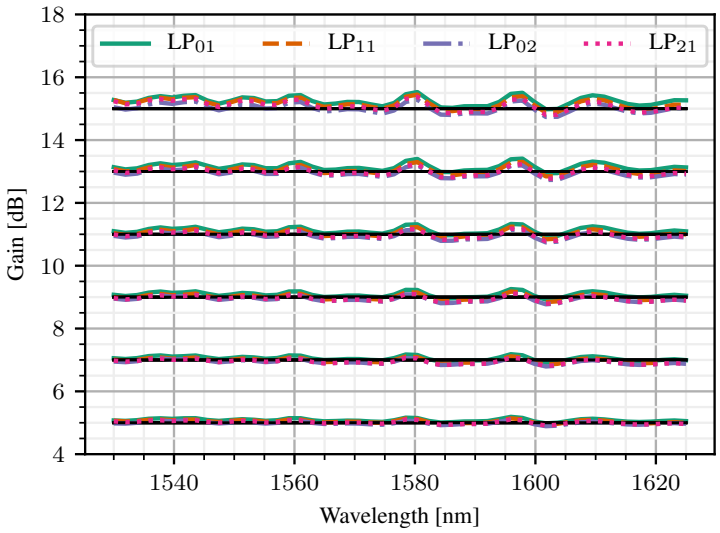


Figure 2.9: Target and predicted flat gain profiles for a 4-mode fiber over the C+L–band, using 8 pumps.

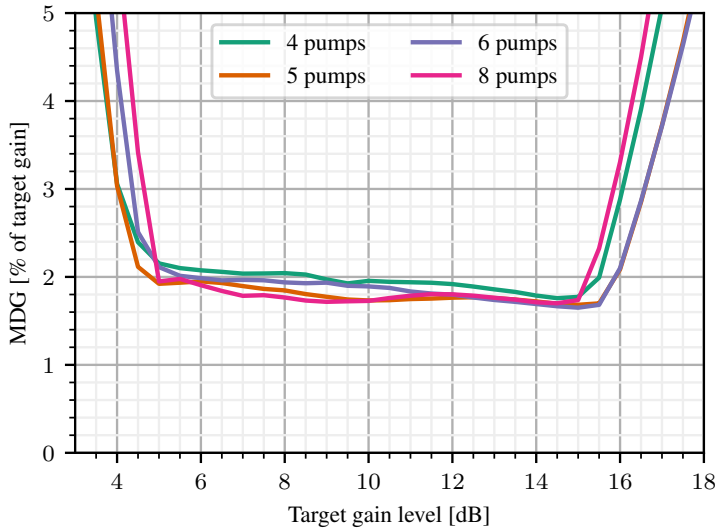


Figure 2.10: Relative MDG as a function of the gain level, for different number of pumps.

by the AE, as a function of the target gain level; the cases of 4, 5, 6, and 8 pumps are considered, with solid lines representing the average power and shaded areas depicting the power variation by employing different numbers of pumps. Independently of the amplification level, no power is launched in the LP_{01} and LP_{11} modes, with 70% of the total power assigned to LP_{21}, and the remaining 30% to LP_{02}, confirming the results of [56] and [46]. Even though no power is injected in LP_{01} and LP_{11}, these two modes are those that experience the highest amplification, predominantly contributing to the residual MDG of

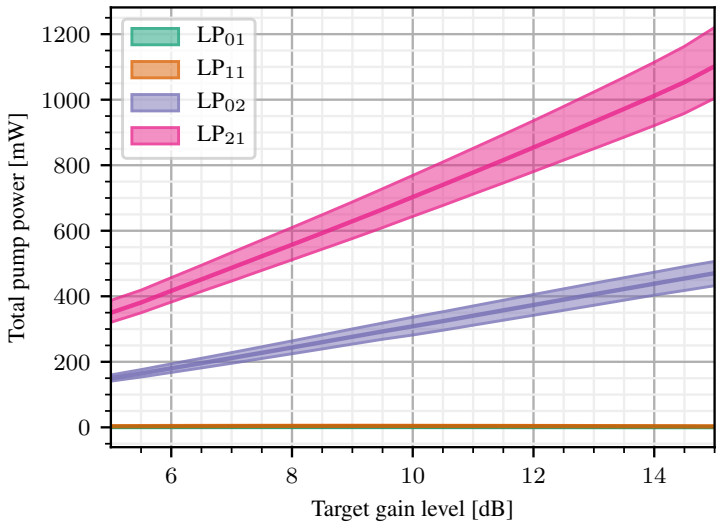


Figure 2.11: Total pump power at $z = L$ in each mode of FMF1, as a function of the target gain level. Shaded areas indicate the variation using different number of pumps.

the system. In order to confirm the role of the overlap integrals in determining the MDG we test two additional 4-mode fibers. The first, which we label "FMF2", is a SIF with a core diameter of 18 µm, core refractive index of 1.466, and a relative refractive index difference between core and cladding $\Delta = 0.4\%$, supporting the propagation of the LP_{01}, LP_{11}, LP_{02} and LP_{21} modes over the entire simulation bandwidth. Its overlap integrals are reported in Table 2.2. The second fiber, which we refer to as "FMF3", is instead an ideal 4-mode fiber whose overlap integrals are equal to $5.47 \times 10^{-9}\,\mathrm{m}^{-2}$, i.e the overlap integral for the LP_{01}-LP_{01} mode pair of FMF2. All the other simulation parameters, including the attenuation spectrum and Raman gain coefficient of the fiber, remain unchanged. Training the AE under the same conditions, we can observe the effect of the fiber design on the performance of the system in terms of residual MDG. For the case of 8 Raman pumps, we report the MDG for the three considered fibers as a function of the gain level in Fig. 2.12: FMF2 exhibits the highest MDG among the fibers, reaching a value of approximately 4% of the target gain inside the training interval of 5 dB to 15 dB, while for FMF3 the AE correctly predicts the power distribution among the modes that results in no MDG, launching power only in the LP_{01} mode.

2.3.4 *Tilted gain profiles*

In order to account for tilted gain profiles, the AE is trained using ideal gain profiles with average gain and tilt uniformly sam-

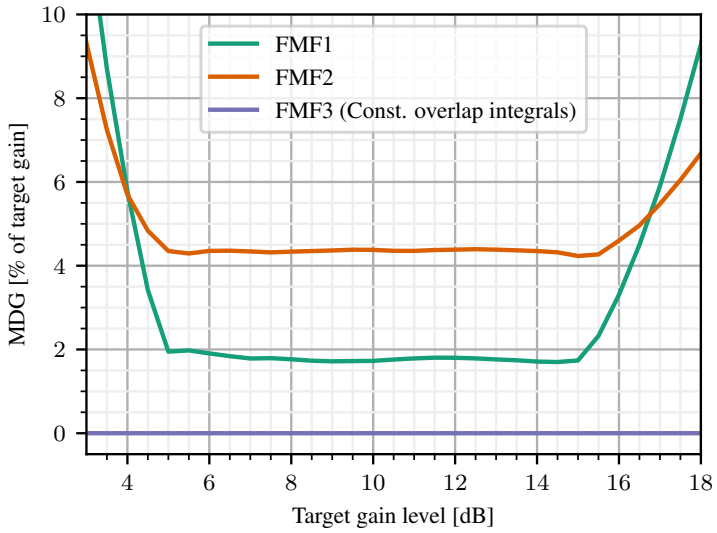


Figure 2.12: Relative MDG for flat target gain profiles, as a function of the gain level, for three different fibers. The number of pumps is set to $N_p = 8$.

pled from the two-dimensional training region $\mathcal{T} = [5, 15]$ dB $\times$ $[-0.015, 0.015]$ dB nm^{-1}, resulting in a maximum total tilt on the C+L band equal to $T_{\text{max}} = 0.015$ dB nm^{-1} $\times$ 95 nm $= 1.425$ dB. In Fig. 2.13 we report the target gain profiles and corresponding AE predictions using FMF1 and 8 pumps, for a total tilt equal to T_{max} and for different average gain levels inside the training region. Results show good agreement between targets and predictions, with approximately the same gain profile on each mode, up to the residual MDG.

An analysis similar to that of flat gain profiles is carried out for the case of tilted profiles, evaluating the metrics of interest for FMF1 and varying the number of employed Raman pumps, keeping the other simulation parameters unchanged. We compute RMSE, flatness, and MDG of the predicted gain profiles and visualize them in Fig. 2.14, representing the metrics as a function of the target gain level and total tilt on the C+L band. Each metric is reported in terms of percentage of the target gain level; for RMSE and flatness we consider the worst-case scenario among the modes, i.e. their maximum value. Fig. 2.14 is organized such that columns 1 through 4 of the grid correspond to the case of 4, 5, 6, and 8 pumps, whereas row 1, 2, and 3 correspond to RMSE, flatness and MDG, respectively. The color scale for each metric is saturated to different levels in order to improve the contrast of the color maps. In Fig. 2.14 (a)–(d) we can appreciate the improvements in terms of RMSE by using more pumps: the color map is increasingly darker inside and in the vicinity of the training region $\mathcal{T}$, whose bounds are represented by a dashed rectangle. Additionally, by using 5 or more pumps, the level curves show that a RMSE lower than 3% of

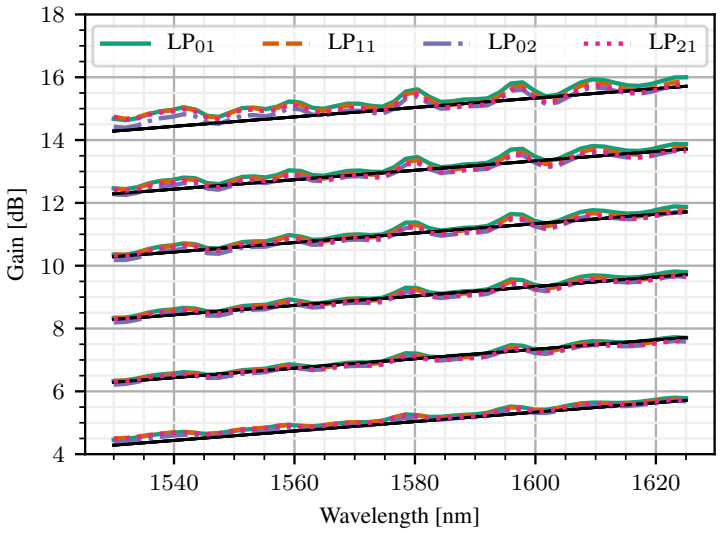


Figure 2.13: Target and predicted gain profiles for the tilted case, using 8 pumps and with a total tilt of 1.425 dB, i.e. the maximum considered tilt during training.

the target gain level is achieved for (practically) all the gain level-tilt combinations in $\mathcal{T}$.

Similar observations can be made for the flatness from Fig. 2.14 (e)–(h), where a value of about 17% is reached for the points inside the training region using 4 pumps; increasing the number of pumps leads to progressively lower flatness values, down to 5% inside $\mathcal{T}$ with 8 pumps.

Similarly, for the MDG, Fig. 2.14, (i)–(l) show that a higher number of pumps brings no significant changes, as the minimum achievable MDG is determined by overlap integrals of the fiber. Its value stays infact approximately constant inside the training region regardless of the pump count, with the level curve showing that MDG values lower than 4% are achieved for a region considerably wider than $\mathcal{T}$.

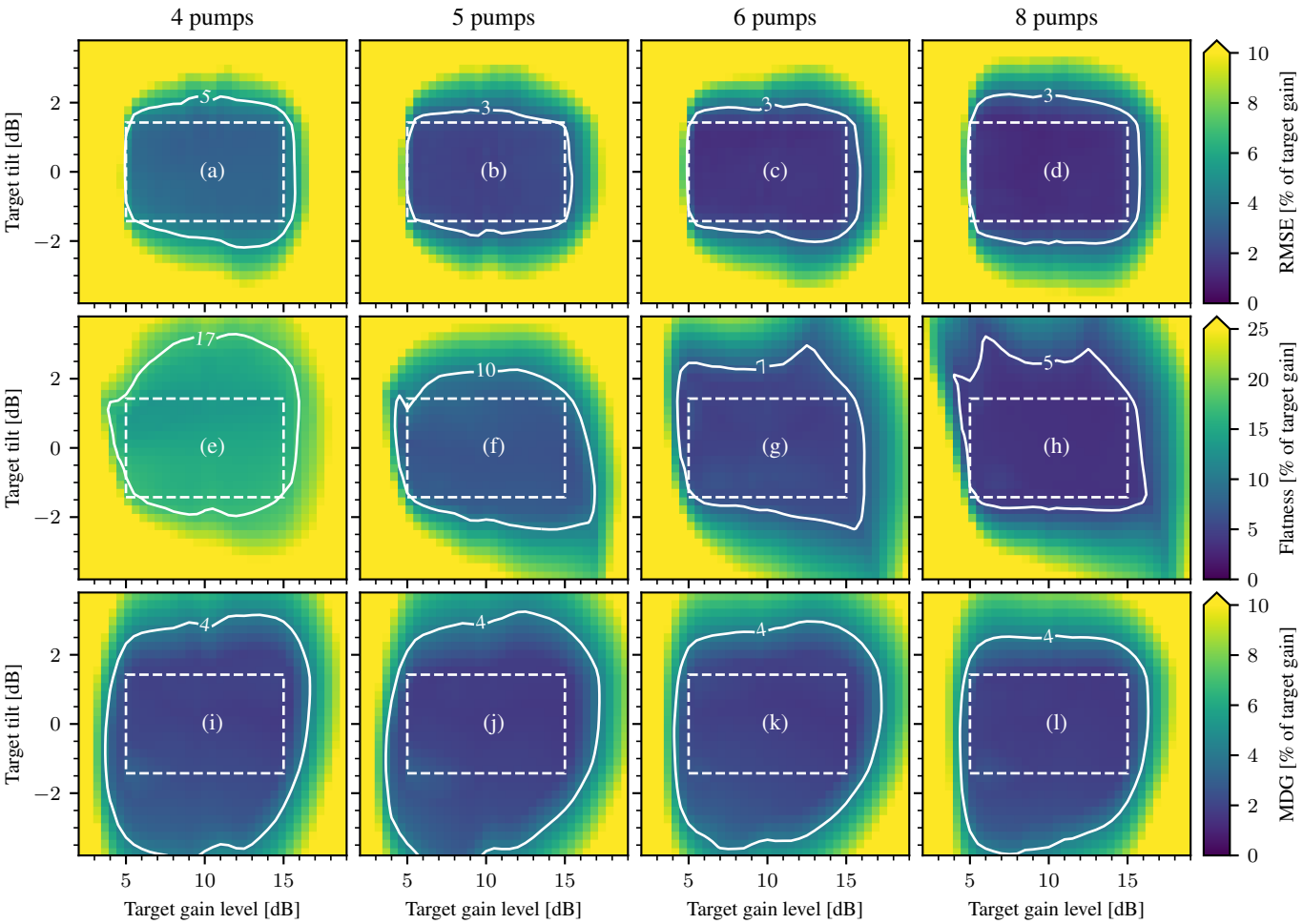


Figure 2.14: Calculated metrics for the tilted gain case, varying the number of Raman pumps: RMSE (a)–(d), flatness (e)–(h), and MDG (i)–(l) as a function of the target gain level and target tilt. For RMSE and flatness their maximum value among the modes is reported. Columns 1 through 4 refer to the case of 4, 5, 6, and 8 pumps, respectively.

2.4 NOISE CHARACTERISTICS AND SENSITIVITY TO POWER VARIATION

As introduced in chapter 1 and section 2.2, one of the main benefits of using distributed Raman amplification schemes is the improved noise properties with respect to lumped optical amplifiers. Additionally, with the advent of dynamic optical networks and advanced routing and switching algorithms for MDM, accounting for changes in the power of the information channels, especially in multi-span settings, is necessary [74–76]. For this reason, in this section we assess the robustness of the method presented in section 2.3 in terms of the sensitivity of the predicted gain profiles to changes in the power of

the signals to be amplified, and characterize the resulting noise figure (NF) by evaluating its spectral and spatial dependence. The scalability of the approach is demonstrated by considering a SIF supporting the propagation of 6 groups of linearly polarized modes, requiring only the computation of the overlap integrals of eq. (2.2) and changing the size of the input and output layers of the NN to account for the new data dimensionality prior to its training procedure.

2.4.1 *Propagation model*

In FMF RAs, the evolution of the signals, pumps in the presence of amplified spontaneous emission (ASE) is governed by the following nonlinear differential equations [46, 77]:

$$\xi_i \frac{dP_i^m}{dz} = -\alpha_i P_i^m + P_i^m \sum_{j \neq i,n} I_{m,n} g_R(\nu_i, \nu_j) P_j^n, \tag{2.29}$$

$$\begin{aligned} \xi_i \frac{dN_i^m}{dz} = &-\alpha_i N_i^m + N_i^m \sum_{j \neq i,n} I_{m,n} g_R(\nu_i, \nu_j) P_j^n \\ &+ 2h\nu_i B_{ref} \sum_{j \neq i,n} \eta_{i,j} I_{m,n} g_R(\nu_i, \nu_j) P_j^n, \end{aligned} \tag{2.30}$$

where P_i^m and N_i^m are the signal/pump and ASE power at the ith frequency and mth mode, with $i = 1, \ldots, N_s + N_p$, $m = 1, \ldots, M$, and N_s, N_p, M the number of signals, pumps and modes, respectively; ξ_i is equal to -1 at the frequency of counterpropagating pumps and $+1$ otherwise; α_i is the attenuation coefficient at the ith frequency; $g_R(\nu_i, \nu_j)$ is the Raman gain coefficient between frequencies ν_i and ν_j. Finally, $I_{m,n}$ are the overlap integrals between mode m and n defined in eq. (2.2). The noise source term $\eta_{i,j}$ is the phonon occupancy factor between the ith and jth frequencies, defined as [21, 77]

$$\eta_{i,j} = 1 + \frac{1}{\exp\left[\frac{h|\nu_i - \nu_j|}{k_B T}\right] - 1}, \tag{2.31}$$

where h is the Planck constant, k_B is the Boltzmann constant, T is the temperature of the fiber and B_{ref} is the reference bandwidth over which noise is computed for a specific wavelength. For signal frequencies $i = 1, \ldots, N_s$ and mode m the on-off gain is again defined as in eq. (2.3).

$I_{m,n}$	LP_{01}	LP_{11}	LP_{21}	LP_{02}	LP_{31}	LP_{12}
LP_{01}	4.42	2.97	2.14	3.56	1.58	2.84
LP_{11}	2.97	4.76	2.87	1.83	2.44	2.84
LP_{21}	2.14	2.87	4.48	1.69	2.8	1.43
LP_{02}	3.56	1.83	1.69	5.12	1.73	2.48
LP_{31}	1.58	2.44	2.8	1.73	4.25	1.3
LP_{12}	2.84	2.84	1.43	2.48	1.3	4.2

Table 2.3: Values of $I_{m,n}$ as defined in Eq. (2.2) (units of $10^9\,m^{-2}$).

2.4.2 *Results*

The considered fiber is a $L = 70$ km long SIF supporting 6 groups of LP modes, and whose computed overlap integrals are reported in Tab. 2.3. The fiber attenuation coefficient is again described by a polynomial fitted to attenuation data of a commercial fiber, illustrated in fig. 2.6.

The Raman response function that has been utilized in the previous section, and detailed in [59] is employed with a peak Raman gain coefficient $g_R = 7 \times 10^{-14}\,W^{-1}\,m$. The NN is a feed-forward NN with 5 fully-connected layers of 1000 hidden neurons each, rectified linear unit activations, and a sigmoid function applied to the output to scale the values predicted by the NN to actual wavelengths and powers in a specified range.

The resulting NN is trained using the RMSE cost function and the Adam algorithm, considering input signals distributed on a 50 wavelengths WDM comb on the C+L band, with an input power of -20 dBm per channel, while 6 counterpropagating pumps are employed.

The training algorithm detailed in the previous section is run for 1000 iterations, using batches of 1024 curves with average gain and tilt uniformly sampled from the intervals from 5 to 15 dB, and from -0.015 to 0.015 dB/nm, respectively.

In Fig. 2.15 (a) and (b) the excellent results in terms of mean error and flatness are confirmed, for the case of flat and tilted gain profiles, respectively, obtaining results that are comparable to those obtained with 4-mode fibers presented in the previous section. Fig. 2.15 (c) illustrates a close-up view of the predicted gain for the target gain level of 11 dB, showing that the LP_{11} mode is the most amplified, followed by the LP_{02}, and then the rest of the modes with MDG close to zero.

The robustness of the predictions is tested by evaluating flatness, MDG, and mean error with respect to flat target profiles, varying the

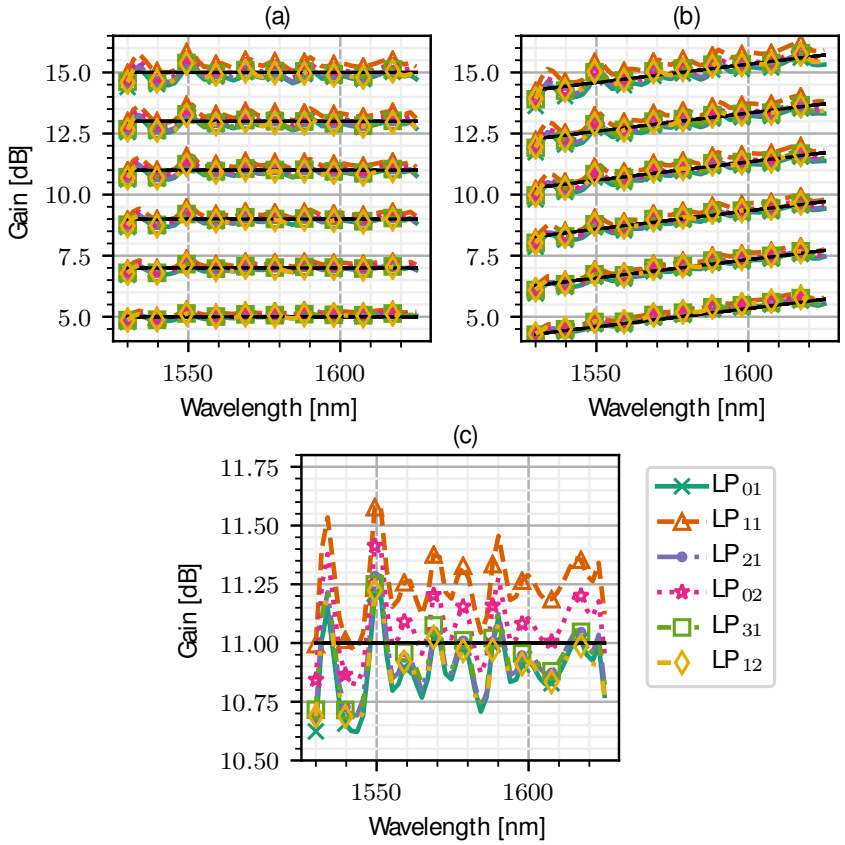


Figure 2.15: (a) Predicted gain for the flat case. (b) Predictions for the tilted case. (c) Close-up view of the 11 dB gain profile.

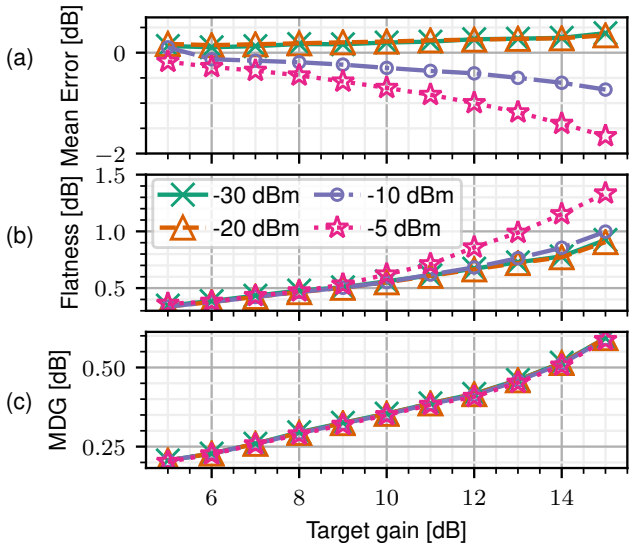


Figure 2.16: Mean error (a), flatness (b) and MDG (c) for flat target gains, varying the signal input power.

input power of the transmitted signals. Mean error and flatness are computed as their maximum value among the different modes.

The mean error is reported in Fig. 2.16 (a), showing that increasing the signal power leads to under-amplification at higher gains. This is clearly visible in Fig. 2.17, where the predicted gain curves for each corresponding target are plotted for different gain levels when the input signal power is either -5 dBm (a) or -10 dBm (b). The cause of this behavior is found in the depletion of the Raman pumps that starts occurring with increasing power. In fact, since the employed NN is trained at low input signal power, the pump depletion is not captured by the resulting model, which exhibits degraded generalization properties when this effect is not negligible. A potential approach to learn the pump depletion dynamics and improve the model generalization without training multiple NNs could consist in adding the input signal power level as an additional input of the NN. By associating each gain profile in the training batches to a different input power value sampled from a selected interval, e.g. from -30 dBm to -5 dBm, the NN can, in principle, learn to include pump depletion when the input power is sufficiently high.

From Fig. 2.16 (b), similar comments can be made about flatness, which rises to over 1 dB when using -5 dBm per channel. For the other cases, the gain flatness is approximately 6% of the total gain. On the contrary, varying the input signal power has no significant effect on the MDG, which remains equal to about 3.5% of the target gain level (Fig. 2.16 (b)).

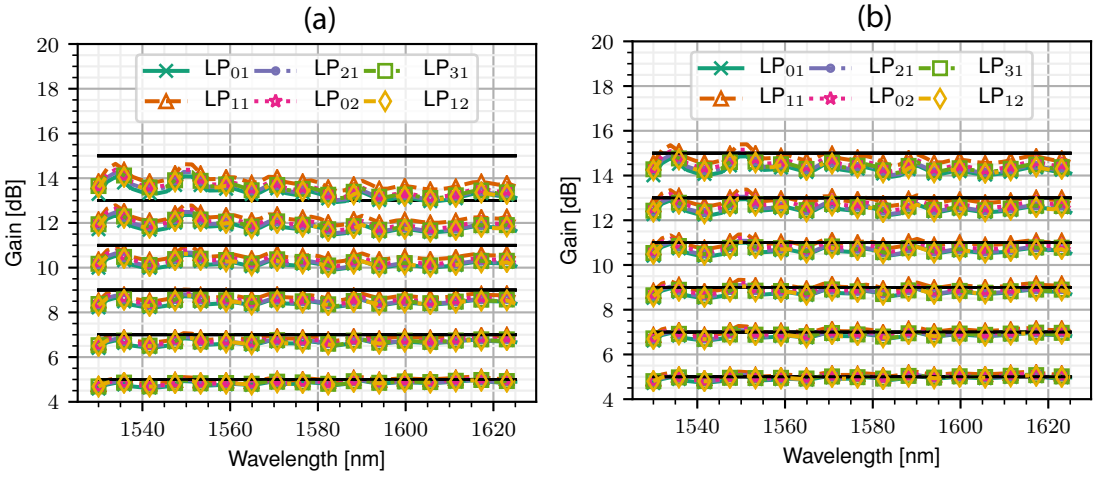


Figure 2.17: Predicted gain for the flat case when the input signal power is (a) -5 dBm and (b) -10 dBm

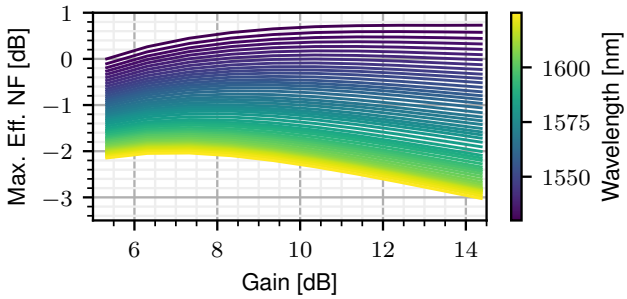


Figure 2.18: Maximum effective noise figure among modes as a function of the gain, for every signal wavelength.

Regarding the analysis of the noise characteristics of the obtained distributed amplifier, the output ASE spectrum is computed by solving Eq. (2.30) with the predicted pump parameters and used to determine the effective NF at the ith frequency and mth mode [21]:

$$\mathrm{NF}_{\mathrm{eff}}(i,m) = \frac{\mathrm{OSNR}_{in}}{\mathrm{OSNR}_{out}} \cdot \frac{1}{\mathcal{L}(\nu_i)} = \frac{P_i^m(0)}{h\nu_i B_{ref}} \cdot \frac{N_i^m(L)}{P_i^m(L)} \cdot \frac{1}{\mathcal{L}(\nu_i)}, \qquad (2.32)$$

where $\mathcal{L}(\nu_i)$ are the total link losses at frequency ν_i. In Fig. 2.18 we report the maximum $\mathrm{NF}_{\mathrm{eff}}$ among the modes as a function of the achieved gain, for each signal wavelength, using a reference noise bandwidth of 0.5 nm. Recalling that the effective NF of a distributed amplifier describes the NF that a lumped amplifier would need to match its performance [21], fig. 2.18 shows that this RA always outperforms any quantum-limited lumped amplifier, whose noise figure can be at most 3 dB. The $\mathrm{NF}_{\mathrm{eff}}$ difference among modes is then computed varying the power of the input signals, using the pump parameters

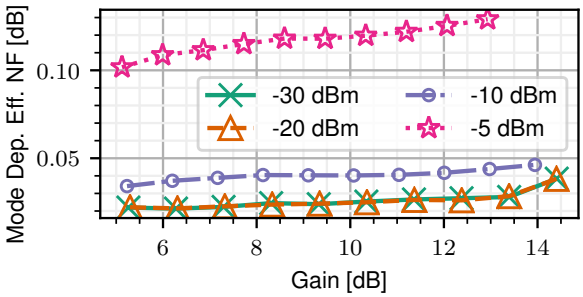


Figure 2.19: The mode-dependent effective noise figure as a function of the Raman gain, for various signal power levels.

predicted by the model trained with -20 dBm per channel. These results are reported in Fig. 2.19, showing that for low input signal power NF_{eff} is mode-equalized, with the total variation remaining below 0.05 dB for every gain level, while rising to about 0.1 dB when -5 dBm per channel are injected in the fiber.

2.5 SIMULTANEOUS GAIN AND OSNR OPTIMIZATION

While the noise figure is essentially perfectly equalized over the spatial channels, its variation over the considered bandwidth is substantial, as can be seen from Fig. 2.18. In fact, the for higher gain levels, the noise figure of the distributed amplifier shows a variation of almost 4 dB. This is a known issue that is also present in ultrawide band SMF Raman amplifiers, caused mainly by the fact that all the Raman pumps are employed in a counterpropagating scheme [21].

In fact, in all-counterpumping schemes, the interactions occurring between pumps cause the shorter wavelength pumps to deplete more quickly compared to longer wavelength. As a consequence, the penetration depth of shorter wavelength pumps is significantly reduced, providing little distributed amplification to the short wavelength signals. On the contrary, long wavelength pumps are amplified and are thus able to penetrate deeper in the span, providing more distributed amplification to long wavelength signals, and consequently improve their noise figure, as seen in Fig. 2.18. Moreover, the same interaction between pumps occurs between signals as well, causing the power transfer from short to long wavelengths, especially in longer spans. As a result, short signal wavelengths experience additional loss, resulting in a degraded OSNR. A solution to overcome this issue, and compensate the resulting OSNR tilt, is to preferentially copump the shorter wavelengths and optimize the pump parameters accordingly [78, 79].

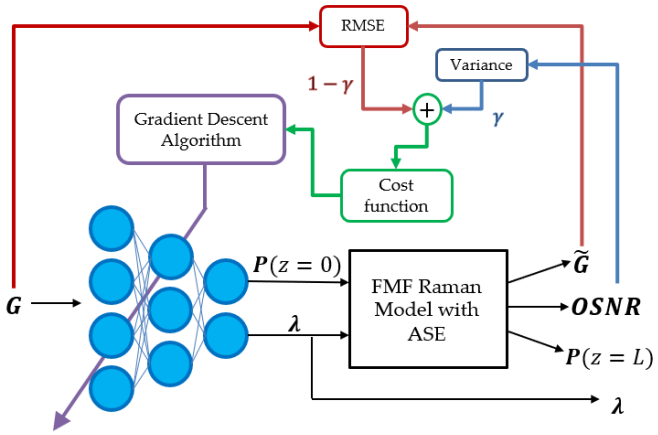


Figure 2.20: Diagram of the ML model for simultaneous gain and OSNR optimization. The computation of the cost function to minimize is also displayed.

2.5.1 *Proposed method*

In order to test this approach also in the framework of FMF Raman amplifiers, the AE that was presented in the previous section has been modified to support bidirectional pumping schemes and to simultaneously optimize both the gain profile and the OSNR spectrum, with the objective of compensating the tilt resulting from fully-conterpumping amplifiers.

Specifically, the decoder network $\mathcal{R}$ implementing the differentiable Raman model has been extended with the inclusion of eq. (2.30), describing the evolution of ASE noise, in addition of the usual equations governing the interaction between pumps and signals. Owing to this addition, while training the model, the backpropagation algorithm is able to compute the gradients of the NN parameters also with respect to quantities depending on ASE power, meaning that the NN can be optimized for e.g. the OSNR spectrum.

In case of bi-directional pumping schemes, the output of the encoder NN $\mathcal{N}$ needs to be interpreted accordingly. In particular, the pump powers predicted by the encoder are all assumed at position $z = 0$, i.e. at the start of the fiber. This means that for co-propagating pumps the NN directly predicts the power to inject at a specific wavelength/mode combination, while for counterpropagating pumps, the NN predicts the power of the pump after it has propagated for the entire span. The actual power to inject for counterpropagating pumps at position $z = L$, i.e. at the end of the span, is given as the output of the decoder, after the integration of the propagation equations, allowing us to avoid

the use of shooting algorithms, as described in the previous section. Additionally, at the output of decoder network the OSNR spectrum is also computed for the specific combinations of pump parameters predicted by the encoder, giving the following relations for the ML model under consideration

$$\tilde{\lambda}_k, \tilde{P}_k^n(z=0) = \mathcal{N}(G_i^m) \tag{2.33}$$

$$\tilde{G}_i^m, \tilde{P}_k^n(z=L), O_i^m = \mathcal{R}(\tilde{\lambda}_k, \tilde{P}_k^n(0)), \tag{2.34}$$

where $\tilde{\lambda}_k$, $\tilde{P}_k$ are the predicted pump wavelength and power for wavelength k and mode m, while $\tilde{G}_i^m$, G_i^m, O_i^m the predicted gain profile, the target gain profiles, and the corresponding OSNR at wavelength k and mode m.

With these additions to the ML model, we can now train the encoder NN to learn the correct pump parameters to generate a given target profile and simultaneously flatten the output OSNR spectrum for each mode.

To do so, it is also necessary to update the cost function that the training algorithm tries to minimize. Here we consider a cost function that is made up of two contributions: the first contribution is unchanged, and is the RMSE between target and predicted gain profiles; the second instead must be proportional to the OSNR variation. A differentiable and easy to compute metric is given by its variance over the considered spectrum, averaged over the spatial channels. The two terms are weighted by a design parameter $\gamma \in [0, 1]$, obtaining a cost function for each target gain profile given by the following

$$\mathcal{C}(\mathbf{G}, \tilde{\mathbf{G}}, \mathbf{O}; \gamma) = (1-\gamma)\frac{1}{M}\sum_{m=1}^{M} \underset{i}{\mathrm{RMSE}}\left(G_i^m, \tilde{G}_i^m\right) \tag{2.35}$$

$$+ \gamma \frac{1}{M}\sum_{m=1}^{M} \underset{i}{\mathrm{Var}}(O_i^m) \tag{2.36}$$

The total cost function is computed as the average of eq. (2.35) over the gain profiles contained the batch of data in each iteration of the training algorithm.

The newly defined model, along with the computational steps to compute the cost function, are illustrated in fig. 2.20.

2.5.2 *Results*

The newly defined ML model has been tested on a 70 km long 4-mode SIF with a core diameter of 18 µm, core refractive index of 1.466, and a relative refractive index difference between core and cladding

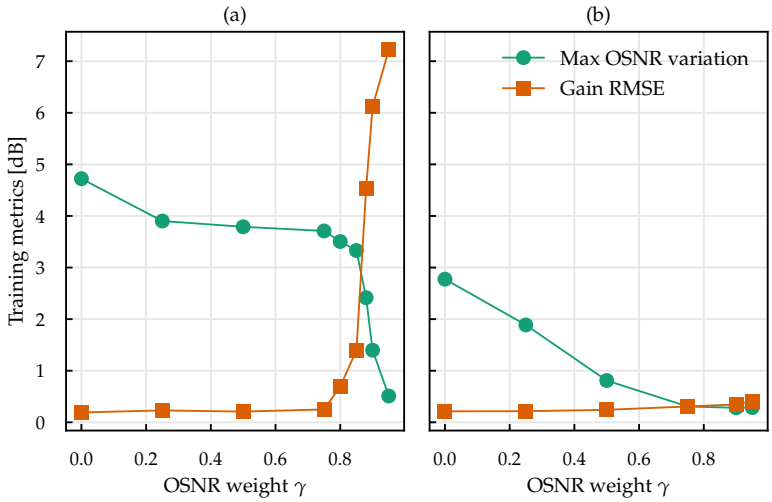


Figure 2.21: Maximum OSNR variation and gain RMSE for the trained model, as a function of the parameter γ, for the fully counterpumping case (a) and bidirectional pumping (b).

$\Delta = 0.4\%$, supporting the propagation of the LP_{01}, LP_{11}, LP_{02} and LP_{21} modes over the entire simulation bandwidth. The encoder NN is a FC NN with 5 hidden layers, consisting of 1000 neurons each, using ReLU activation functions between each layer, and with a sigmoidal activation at the output layer, enabling us to set the initial condition for the pump parameters as described in the previous section. The bandwidth dedicated to the pump remains unchanged with respect to the previously tested models, as well as the number of WDM channels (50), their bandwidth (C+L) and their initial power (−20 dBm).

We test the effect of the weight parameter γ in two scenarios: the first where all the 8 Raman pumps are counter-propagating, and the second one in which the 4 pumps at shorter wavelengths are co-propagating with the signal. Each NN is trained with the Adam algorithm running for 1000 iterations, with batches of 512 random tilted profiles with average gain ranging from 5 dB to 15 dB; using the technique previously detailed, we lock the updating of the pump wavelengths for the first 100 iterations, facilitating the optimization of the pump power first, which is critical for the convergence of the algorithm.

In fig. 2.21 (a) we report both the gain RMSE and the total variation of OSNR at the end of the last training iteration for different values of the weighting parameter γ in the interval $(0, 1)$. For moderate values of γ the OSNR tilt improvement is very modest, going from almost 5 dB to slightly below 4 dB, with the gain RMSE being essentially unaffected. When γ reaches the critical value of approximately 0.8

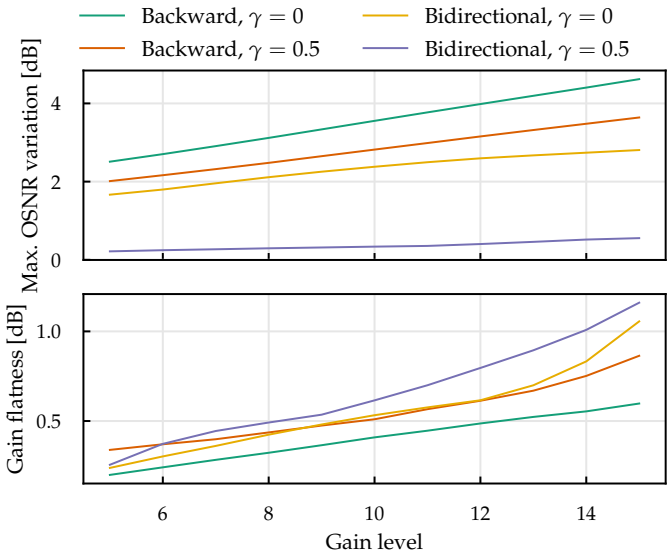


Figure 2.22: Maximum OSNR variation (top) and gain flatness (bottom) over the considered spectrum, for different values of average gain levels, using different pumping schemes and values of the parameter γ.

we instead see a dramatic improvement for the OSNR, which reaches below 0 dB, but at the expense of the accuracy on the predicted gain, which instead quickly diverges to over 7 dB, confirming that the OSNR optimization is not achievable without changing the pumping scheme.

The results for the bidirectional pumping scheme are instead reported in fig. 2.21 (b). In this case, it is apparent how a significant improvement of the OSNR is possible without seriously compromising the error on the predicted gain. With increasing γ, the total OSNR tilt decreasing from 3 dB when $\gamma = 0$, to approximately 0.3 dB when $\gamma = 1$, while at the same time the total RMSE on the predicted gain profiles increases from 0.2 dB to about 0.4 dB. A good trade-off is obtained by weighting the two contributions equally by setting $\gamma = 0.5$, for which the maximum OSNR variation decreases below 1 dB with an increase in RMSE of 0.4 dB.

Now that a suitable value of γ is identified, we can evaluate how the different schemes behave for the case of flat target profiles. In fig. 2.22 we visualize the gain flatness (bottom) and OSNR variation (top) as a function of different levels of average gain. We compare the counterpumping and bidirectional pumping case when the NN is trained to minimize only the gain error ($\gamma = 0$) and when a simultaneous optimization of gain and OSNR is demanded ($\gamma = 0.5$). All the metrics reported consider the worst-case among the different spatial channels, i.e. are computed as the maximum over the modes.

As expected, in terms of OSNR variation the worse case is given by the counterpropagating setup without OSNR optimization, exhibiting variations of over 4 dB at high amplification levels. When including the OSNR in the cost function with the value of $\gamma = 0.5$, a slight improvement is obtained, with a maximum decrease of 1 dB when a gain of 15 dB is required. If considering a bidirectional pumping scheme, a similar improvement is obtained, even when no optimization of the OSNR is explicitly requested by adding it to the cost function. Finally, using this scheme with $\gamma = 0.5$ we achieve the biggest improvement, with the worst-case tilt, obtained at high-gain levels, reaching just over 0.5 dB.

Regarding the gain flatness instead the order is the opposite, with the optimal configuration given by a fully counterpumping amplifier with $\gamma = 0$, and with an OSNR-optimized bidirectional scheme yielding the highest flatness values of about 6 percent of the average gain. This behavior has two explanations. First, and maybe obvious observation, is that by reducing its contribution to the total cost function, is it expected to obtain a higher error on the gain prediction under the same pumping conditions. Secondly, the worse flatness value of the bidirectional scheme compared to the counterpropagating one, when the total number of pumps is the same, is also explained by the fact that in bidirectionally-pumped amplifiers, co- and counterpropagating pumps only partially interact, especially in longer spans, virtually reducing the degrees of freedom in the optimization. This has been shown also in [21], where the same flatness achieved in the counterpumping scheme can only be reached with two additional pump lasers in the bidirectional configuration.

An example of the predicted profiles when flat target gains are given as the input to the NN trained using the bi-directional scheme and with a weighting parameter $\gamma = 0.5$ is given in fig. 2.23 (left), showing MDG values of approximately 4 percent of the average gain, in line with the results presented so far in this thesis.

Finally, in fig. 2.23 (right) instead, we highlight the improvement in the OSNR using this scheme with respect to the non-optimized, fully-counterpumping case, for the reference gain level of 10 dB, going from approximately 3.5 dB to less than 0.5 dB of total tilt.

2.6 CONCLUSIONS

In this chapter, the use of a machine learning method that leverages a neural network to learn the complex relationship between the gain of a Raman amplifier and the corresponding values of pump powers and wavelengths, first proposed in the context of single-mode fibers, has been applied to the case of few-mode fibers for space-

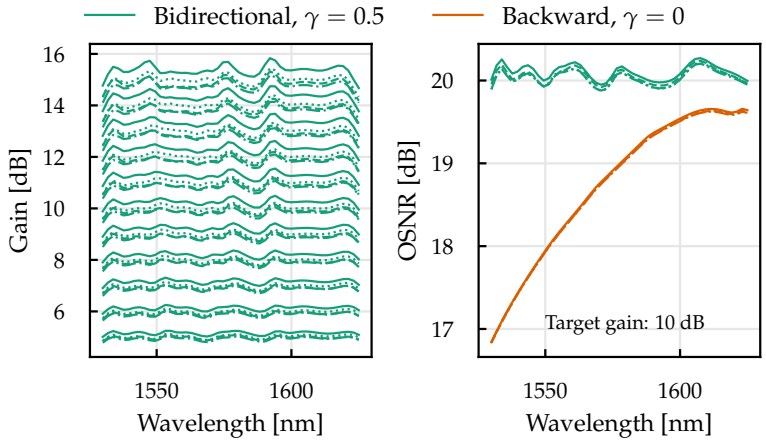


Figure 2.23: Predicted flat gain profiles for the bi-directional pumping scheme with $\gamma = 0.5$ (left), and comparison of the OSNR profiles (right) between the worst-case and best-case scenarios, for an average target gain of 10 dB.

division-multiplexing. The neural network is used to obtain the pump parameters to shape the amplifier gain spectrum and approximate a given target profile. The presented results show that careful design choices in the dataset generation phase can lead to improved accuracy of the predicted Raman gain profiles for the more practical case in which flat or tilted profiles are needed. The proposed techniques have been applied to a distributed Raman amplifier using a few-mode fiber supporting the LP_{01} and LP_{11} mode groups and using 8 pumps, obtaining promising results for a wide interval of amplification levels, obtaining root-mean-square error, gain flatness, and mode-dependent gain values of approximately 0.25 dB, 0.75 dB, and 0.2 dB, respectively, extending the results published in the literature to the broader C+L optical band.

Due to the difficulty of applying this approach to fiber with a higher mode-count, an improved unsupervised machine learning method based on autoencoders has later been proposed. Thanks to automatic differentiation, a numerical Raman model is embedded in the autoencoder, allowing to train it directly on ideal gain profiles (e.g. flat or tilted) and obtaining a robust unsupervised learning method that does not rely on a pre-computed dataset to learn the inverse model. In fact, the relationship between input target gain and the pump parameters that best approximate it are learned in the training phase from the embedded numerical model, allowing to accurately sample the targeted region of the space of possible gain profiles. As a result, this method scales well with respect to the number of fiber modes, the number of Raman pumps, and the amplification bandwidth. On this regard, the

low root-mean-square error (quantified for various number of pump wavelengths) demonstrated the achievement of the target profile.

This approach is tested on a 4-mode fiber using the counter-pumping scheme, various numbers of pumps, up to 8, and for the C+L band. The training process is further simplified by the fact that the autoencoder can directly predict the pump powers at $z = 0$, eliminating the need to employ costly shooting algorithms that are typically needed for counter-propagating Raman amplification models. The pump power to be injected in the fiber are in fact computed with a single integration of the propagation equations. Very good results have been achieved regarding flatness and mode-dependent gain over the entire C+L band and the considered interval of gain levels and tilts, reaching a gain flatness of 3% of the total gain using 8 pumps, and a residual mode-dependent gain of 2% of the total gain, independently of the number of Raman pumps.

The same method has later been applied to a fiber supporting the propagation of 6 groups of LP modes, showing minimal degradation of the aforementioned metrics with respect to the 4-mode case. Additionally, the robustness of the trained models to the variation of the input power of the transmitted signals has been verified, showing that power variations in the order of 10 dB do not significantly affect the quality of the resulting gain spectra. Moreover, the noise characteristics of the generated amplification profiles have been evaluated by solving the equations describing the evolution of amplified spontaneous emission noise. The advantage of distributed amplification schemes compared to lumped amplifiers has been confirmed by showing that their equivalent noise figure is always lower than the quantum-limited 3 dB threshold on the entire bandwidth, and for each tested signal power value. Furthermore, it has been shown that the minimized mode-dependent gain generated by the neural networks results in a negligible variation of the amplifier noise figure over the spatial channels, reaching a maximum of 0.1 dB when the input signal power is more than 20 dB higher than the value used in the training process.

Finally, the method has been modified to include the amplified spontaneous emission equations in the embedded propagation model, showing that by using a bi-directional pumping scheme in conjunction with a modified cost function for the neural network training can lead to a system that is able to predict the pump parameters to simultaneously approximate the required target gain and spectrally equalize the optical signal-to-noise ratio spectrum.

3

EFFECTS OF STRESS BIREFRINGENCE AND CORE ELLIPTICITY IN RAMAN-AMPLIFIED FEW-MODE FIBER LINKS

3.1 INTRODUCTION

One of the major issues arising in space-division multiplexing (SDM) systems is caused by coupling of light between the different spatial channels during propagation [10].

Extensive research has been done to model and characterize the effect of coupling in spatially-multiplexed transmissions, both for multi-core fibers (MCFs) and few-mode fibers (FMFs), defining two main regimes of propagation [80–83].

The first regime, in which mode coupling is strong, relies on massive multiple-input multiple-output (MIMO) digital signal processing (DSP) techniques to equalize the channel distortions and undo the mixing, arguing that the high-degree of coupling between the modes reduces the group delay spread of the fiber, narrowing its impulse response, and consequently lowering the complexity of MIMO equalizers. A strong mixing between spatial channels can be stimulated by employing either coupled-core multi-core fibers (CC-MCFs), in which the distance between cores is made small so that the cores become coupled [84–86], or FMFs with carefully-designed index profiles [33, 87]. The additional benefit of strongly-coupled SDM transmissions is found in the reduction of the variations of mode-dependent loss (MDL) and mode-dependent gain (MDG), maximizing the channel capacity [88]. Several long-haul experiments have been carried out using CC-MCFs, demonstrating the reliability of MIMO DSP on transmission distances of several thousands of kilometers [89–92].

The second regime is instead characterized by a low degree of coupling between spatial channels, aiming at reducing or outright removing the need of MIMO techniques, using either fibers that support orbital-angular-momentum (OAM) modes [93, 94] or weakly-coupled multi-core fibers (WC-MCFs) [95–97] in which the core-to-core distance is substantially higher than the mode-field diameter of the fundamental mode of each core, and where each core can be treated as an independent waveguide [98]. This approach is however limited for short-distance applications, as fiber imperfections and external perturbations can induce random crosstalk, which can lead to transmissions

in the strong coupling regime when accumulating with transmission distance [98].

3.1.1 *Modeling Nonlinear effects*

Although SDM has emerged as the solution to surpass the nonlinear Shannon limit of single-mode fibers (SMFs) through the lower power-density enabled by physically separating multiple information streams, nonlinear effects remain the ultimate factor determining the maximum achievable capacity of optical fiber communications. The modeling of such effects is then of paramount importance in the understanding of propagation limits in both FMFs and MCFs. In principle, the interaction between all possible combinations of modes/cores and wavelengths must be studied to completely describe the nonlinear phenomena occurring in SDM transmissions, as was introduced in the previous chapters regarding Raman amplification.

The most general description of such interaction in multimode fibers (MMFs) is given by the Multimode Nonlinear Schrödinger Equation (MM-NLSE) [99, 100], in which the nonlinear interaction is described by a set of $(2N)^4$ overlap integrals between the mode functions, where N is the total number of spatial modes, and the factor 2 accounts for polarization degeneracy.

In the regime of strong mode coupling, it is argued that the linear mode mixing effects occur on length scales much shorter than the nonlinear interaction length [101]. These considerations justify the simplification of the equations by accounting for the nonlinear effects with just a single coefficient, obtained by averaging the overlap integrals appearing in the MM-NLSE, resulting in the generalized Manakov equation [101], which extends the well-known Manakov-PMD equation used in the study of polarization coupling in SMFs [102]. An equivalent equations has also been derived in [103] and extended to the case of MCFs.

A further version of the generalized Manakov equation is obtained when considering strong coupling occurring only between modes belonging to the same group of quasi-degenerate modes of a FMF, with no interaction occurring between different groups. In this case, the propagation of the mode groups are described by the coupled multi component Manakov equations [104], extending the results of [101] by separating the intra-group and inter-group nonlinear effects with two separate averaged nonlinear coefficients.

Using these models, it has been shown that the regime of strong coupling exhibits beneficial properties for the reduction of nonlinear penalties, with [105] and [106] demonstrating that Kerr-induced effects are inversely proportional to the number of cores of a CC-MCF or

modes of a FMF, respectively. Similar results have also been shown in the study of intermodal four-wave mixing (IM-FWM) in FMFs, with [107] demonstrating that suppression factors greater than 40 dB are achievable when strong mode coupling is considered, and that in general the four-wave mixing (FWM) efficiency is always reduced compared to the case in which mode coupling is not present.

For the case of Raman amplification instead, literature is still not as mature, with only [57] extending the coupled multi component Manakov equations with terms describing the Raman interaction under the hypothesis of strong coupling between degenerate modes of the same group, and total absence of linear coupling between groups. Additionally, numerical results showing the effects of mode coupling on Raman amplifications are lacking, regardless of the regime of coupling considered, with only [108] including Raman scattering effects, but only in the context of Raman induced crosstalk between different channels of a wavelength-division multiplexing (WDM) transmission. Furthermore, the models previously introduced generally describe the coupling process through a fully stochastic approach for mathematical convenience, without considering the different physical mechanisms which cause the coupling process, such as the ellipticity of the fiber core, twisting, or bending, as modeled in [15, 109]. Finally, while the Manakov approach is useful to considerably simplify the MM-NLSE and enable its analytical tractability, it only describes the mode-averaged nonlinear effects, without considering its statistics, which can be important to predict the variability among different fiber realizations or due to time-variant effects. This aspect is particularly important for Raman amplifiers, which exhibit considerable gain fluctuations when polarization mode dispersion (PMD) is present in SMFs [110–113]

In this chapter, following the footprints of [111], which presented a model for SMF Raman amplifiers with PMD effects, a set of continuous-wave equations is derived, describing the evolution of the modes at two different wavelengths acting as pump and signal of a forward-pumping Raman amplifier, modeling linear coupling based on the underlying physical phenomena from which it stems. Through numerical integration of the equations, the combined effect of stress birefringence and core ellipticity on the gain statistics of FMF Raman amplifiers over an ensemble of thousands of fiber realizations are studied for different coupling conditions.

3.2 DERIVATION OF THE MODEL

In silica optical fibers, nonlinear phenomena are related to the third order nonlinear polarization of silica, which can be written as [111]

$$\mathbf{P}^{(3)}(\mathbf{r},t) = \frac{\epsilon_0}{2}\sigma[\mathbf{E}(\mathbf{r},t)\cdot\mathbf{E}(\mathbf{r},t)]\mathbf{E}(\mathbf{r},t) \tag{3.1a}$$

$$+\,\mathbf{E}(\mathbf{r},t)\int_0^\infty \epsilon_0 a(\tau)[\mathbf{E}(\mathbf{r},t-\tau)\cdot\mathbf{E}(\mathbf{r},t-\tau)]\,\mathrm{d}\tau \tag{3.1b}$$

$$+\,\mathbf{E}(\mathbf{r},t)\cdot\int_0^\infty \epsilon_0 b(\tau)[\mathbf{E}(\mathbf{r},t-\tau)\mathbf{E}(\mathbf{r},t-\tau)]\,\mathrm{d}\tau \tag{3.1c}$$

where eq. (3.1a) is the instantaneous Kerr contribution, while eq. (3.1b) and eq. (3.1c) are the parallel and cross-polarized delayed Raman responses.

Restricting ourselves to the case of forward-pumped Raman amplifiers, signal and pump frequencies propagate together. By writing the total electric field and third order nonlinear polarization as

$$\mathbf{E} = \mathrm{Re}\Big\{\mathbf{E}_p e^{-j\omega_p t} + \mathbf{E}_s e^{-j\omega_s t}\Big\}, \tag{3.2}$$

$$\mathbf{P}^{(3)} = \mathrm{Re}\Big\{\mathbf{P}_p e^{-j\omega_p t} + \mathbf{P}_s e^{-j\omega_s t}\Big\}, \tag{3.3}$$

and inserting them in eq. (3.1), the following expression is found [111]

$$\mathbf{P}_\ell^{(3)}(\omega_\ell) = \frac{\epsilon_0}{8}[\sigma + 2\tilde{b}(0)](\mathbf{E}_\ell\cdot\mathbf{E}_\ell)\mathbf{E}_\ell^* \qquad (p_1\mathbf{N}_1) \tag{3.4a}$$

$$+\frac{\epsilon_0}{4}[\sigma + 2\tilde{a}(0) + \tilde{b}(0)](\mathbf{E}_\ell^*\cdot\mathbf{E}_\ell)\mathbf{E}_\ell \qquad (p_2\mathbf{N}_2) \tag{3.4b}$$

$$+\frac{\epsilon_0}{4}[\sigma + 2\tilde{a}(0) + \tilde{b}(\omega_\ell-\omega_f)](\mathbf{E}_f^*\cdot\mathbf{E}_f)\mathbf{E}_\ell \qquad (p_3\mathbf{N}_3) \tag{3.4c}$$

$$+\frac{\epsilon_0}{4}[\sigma + \tilde{b}(0) + \tilde{b}(\omega_\ell-\omega_f)](\mathbf{E}_f\cdot\mathbf{E}_\ell)\mathbf{E}_f^* \qquad (p_4\mathbf{N}_4) \tag{3.4d}$$

$$+\frac{\epsilon_0}{4}[\sigma + 2\tilde{a}(\omega_\ell-\omega_f) + \tilde{b}(0)](\mathbf{E}_f^*\cdot\mathbf{E}_\ell)\mathbf{E}_f \qquad (p_5\mathbf{N}_5) \tag{3.4e}$$

where the subscripts ℓ and f can can either be p or s to indicate pump or signal frequency, ω_p and ω_s, respectively. When $\ell = p$, then $f = s$, and viceversa. Inserting these expressions in the Maxwell equations, $\mathbf{E}_s$ and $\mathbf{E}_p$ are found to satisfy the nonlinear Helmholtz equation at their respective frequencies [111]

$$\nabla^2\mathbf{E}_\ell + \frac{\omega_\ell^2}{c^2}\boldsymbol{\epsilon}_\ell\mathbf{E}_\ell = -\frac{\omega_\ell^2}{\epsilon_0 c^2}\mathbf{P}_\ell \tag{3.5}$$

The extension for the case of a FMF begins by expressing the electric field as the sum of the individual modes supported by the fiber as follows

$$\mathbf{E}_\ell = \sum_{\mu=1}^{M_\ell} \mathbf{F}_{\ell\mu}(x,y) A_{\ell\mu}(z) \exp\left(j\beta_{\ell\mu} z\right) = \sum_{\mu=1}^{M_\ell} \mathbf{F}_{\ell\mu}(x,y) L_{\ell\mu}(z), \tag{3.6}$$

where M_ℓ is the number of polarization and spatial modes propagating at frequency ω_ℓ, $\mathbf{F}_{\ell\mu}(x,y)$ is the mode functions of mode μ at the frequency ω_ℓ, $A_{\ell\mu}$ is its complex amplitude, $\beta_{\ell\mu}$ is its propagation constant, and $L_{\ell\mu}(z) = A_{\ell\mu}(z)\exp\left(j\beta_{\ell\mu} z\right)$. From now on, the spatial dependence of $\mathbf{F}_{\ell\mu}(x,y)$ and $A_{\ell\mu}(z)$ will be omitted for brevity.

After substituting eq. (3.6) in eq. (3.5), we start by calculating the Laplacian of the electric field expressed as the sum of the modes

$$\nabla^2 \mathbf{E}_\ell = \sum_{\mu=1}^{M_\ell} \left[L_{\ell\mu} \nabla_\perp^2 \mathbf{F}_{\ell\mu} + \mathbf{F}_{\ell\mu} \frac{\partial^2 L_{\ell\mu}}{\partial z^2} \right] \tag{3.7}$$

$$= \sum_{\mu=1}^{M_\ell} \left\{ A_{\ell\mu} \nabla_\perp^2 \mathbf{F}_{\ell\mu} + \mathbf{F}_{\ell\mu} \left[\frac{\partial^2 A_{\ell\mu}}{\partial z^2} + 2j\beta_{\ell\mu} \frac{\partial A_{\ell\mu}}{\partial z} - \beta_{\ell\mu}^2 A_{\ell\mu} \right] \right\} e^{j\beta_{\ell\mu} z}, \tag{3.8}$$

where we expressed the Laplacian operator as its transversal and longitudinal (with respect to the direction of propagation) components $\nabla^2 = \nabla_\perp^2 + \partial_z^2$.

We can neglect the second derivative of the complex amplitude by using the slowly-varying envelope approximation [114]

$$\left| \frac{\partial^2 A_{\ell\mu}}{\partial z^2} \right| \ll \beta_{\ell\mu} \left| \frac{\partial A_{\ell\mu}}{\partial z} \right| . \tag{3.9}$$

Using this fact, and collecting $A_{\ell\mu}\exp\left(j\beta_{\ell\mu} z\right)$, the left-hand side of eq. (3.5) becomes

$$\nabla^2 \mathbf{E}_\ell + \frac{\omega_\ell^2}{c^2} \boldsymbol{\epsilon}_\ell \mathbf{E}_\ell = \sum_{\mu=1}^{M_\ell} A_{\ell\mu} \left[\nabla_\perp^2 \mathbf{F}_{\ell\mu} + \left(\frac{\omega_\ell^2}{c^2} \boldsymbol{\epsilon}_\ell - \beta_{\ell\mu}^2 \right) \mathbf{F}_{\ell\mu} \right] e^{j\beta_{\ell\mu} z} \tag{3.10a}$$

$$+ \sum_{\mu=1}^{M_\ell} \mathbf{F}_{\ell\mu} 2j\beta_{\ell\mu} \frac{\partial A_{\ell\mu}}{\partial z} e^{j\beta_{\ell\mu} z} \tag{3.10b}$$

3.2.1 *Inclusion of linear mode coupling*

To include the effect of linear coupling between the modes, a small perturbation to the permittivity tensor is applied

$$\boldsymbol{\epsilon}_\ell = \epsilon_\ell^{(u)} + \Delta\boldsymbol{\epsilon}_\ell, \tag{3.11}$$

where $\epsilon_\ell^{(u)}$ is the original, unperturbed, relative dielectric constant.

The right-hand side of eq. (3.10) then becomes

$$\nabla^2\mathbf{E}_\ell + \frac{\omega_\ell^2}{c^2}\boldsymbol{\epsilon}_\ell\mathbf{E}_\ell = \sum_{\mu=1}^{M_\ell} A_{\ell\mu}\left[\nabla_\perp^2\mathbf{F}_{\ell\mu} + \left(\frac{\omega_\ell^2}{c^2}\epsilon_\ell^{(u)} - \beta_{\ell\mu}^2\right)\mathbf{F}_{\ell\mu}\right] e^{j\beta_{\ell\mu}z} \tag{3.12a}$$

$$+ \sum_{\mu=1}^{M_\ell} \frac{\omega_\ell^2}{c^2}\Delta\boldsymbol{\epsilon}_\ell\mathbf{F}_{\ell\mu}A_{\ell\mu}e^{j\beta_{\ell\mu}z} \tag{3.12b}$$

$$+ \sum_{\mu=1}^{M_\ell} \mathbf{F}_{\ell\mu}2j\beta_{\ell\mu}\frac{\partial A_{\ell\mu}}{\partial z}e^{j\beta_{\ell\mu}z} \tag{3.12c}$$

The term in square brackets of the first term of eq. (3.12) is the Helmholtz equation, which vanishes under the hypothesis of the mode distributions not being affected by nonlinearity, giving

$$\nabla^2\mathbf{E}_\ell + \frac{\omega_\ell^2}{c^2}\boldsymbol{\epsilon}_\ell\mathbf{E}_\ell = \sum_{\mu=1}^{M_\ell}\frac{\omega_\ell^2}{c^2}\Delta\boldsymbol{\epsilon}_\ell\mathbf{F}_{\ell\mu}A_{\ell\mu}e^{j\beta_{\ell\mu}z} + \sum_{\mu=1}^{M_\ell}\mathbf{F}_{\ell\mu}2j\beta_{\ell\mu}\frac{\partial A_{\ell\mu}}{\partial z}e^{j\beta_{\ell\mu}z}. \tag{3.13}$$

We left-multiply by $\mathbf{F}_{\ell\nu}^\dagger$ and integrate on the infinite transverse plane, obtaining

$$\sum_{\mu=1}^{M_\ell}\frac{\omega_\ell^2}{c^2}A_{\ell\mu}e^{j\beta_{\ell\mu}z}\iint\mathbf{F}_{\ell\nu}^\dagger\Delta\boldsymbol{\epsilon}_\ell\mathbf{F}_{\ell\mu}\,\mathrm{d}x\mathrm{d}y + \sum_{\mu=1}^{M_\ell}2j\beta_{\ell\mu}\frac{\partial A_{\ell\mu}}{\partial z}e^{j\beta_{\ell\mu}z}\iint\mathbf{F}_{\ell\nu}^\dagger\mathbf{F}_{\ell\mu}\,\mathrm{d}x\mathrm{d}y. \tag{3.14}$$

In the weakly guiding approximation, the following orthogonality conditions [15, 115]

$$\frac{\beta_{\ell\nu}}{2\omega_\ell\mu_0}\iint\mathbf{F}_{\ell\nu,\perp}^\dagger\mathbf{F}_{\ell\mu,\perp}\,\mathrm{d}x\mathrm{d}y = \delta_{\nu,\mu}\,, \tag{3.15}$$

are used to reduce the sum in the second term of eq. (3.14) to a single term, obtaining

$$\sum_{\mu=1}^{M_\ell} \frac{\omega_\ell^2}{c^2} A_{\ell\mu} e^{j\beta_{\ell\mu} z} \iint_{-\infty}^{+\infty} \mathbf{F}_{\ell\nu}^\dagger \Delta\boldsymbol{\epsilon}_\ell \mathbf{F}_{\ell\mu} \, \mathrm{d}x\mathrm{d}y + 2j\beta_{\ell\nu} \frac{\partial A_{\ell\nu}}{\partial z} e^{j\beta_{\ell\nu} z} \iint_{-\infty}^{+\infty} \mathbf{F}_{\ell\nu}^\dagger \mathbf{F}_{\ell\nu} \, \mathrm{d}x\mathrm{d}y \tag{3.16}$$

Since eq. (3.15) only considers the components of the field perpendicular to the propagation direction, in the derivation of the previous equation we neglect the effect of the longitudinal components, whose magnitude is much weaker in the case of LP modes [15].

The orthogonality conditions in eq. (3.15) also determine the units of the mode functions $\mathbf{F}_{\ell\nu}(x, y)$, which become $\sqrt{\Omega}/\mathrm{m}$. Together with the fact that the units of the total electric field must be $\mathrm{V\,m^{-1}}$, the complex amplitude $A_{\ell\nu}$ takes units of $\mathrm{V}/\sqrt{\Omega}$, meaning that the magnitude squared of the complex amplitudes has units of W.

Indeed, the total power of the electromagnetic field at frequency ω_ℓ can be calculated by integrating the Poyinting vector [115, 116]

$$P_{\mathrm{tot}}(\omega_\ell) = \frac{1}{2} \iint (\mathbf{E}_\ell \times \mathbf{H}_\ell^*) \cdot \hat{\mathbf{z}} \, \mathrm{d}x\mathrm{d}y \tag{3.17}$$

$$= \sum_{\mu=1}^{M_\ell} \iint \frac{\beta_{\ell\mu}}{2\omega_\ell \mu_0} |A_{\ell\mu}|^2 \mathbf{F}_{\ell\mu}^\dagger \mathbf{F}_{\ell\mu} \, \mathrm{d}x\mathrm{d}y \tag{3.18}$$

$$+ \sum_{\nu,\mu,\nu\neq\mu}^{M_\ell} \iint \frac{\beta_{\ell\nu}}{2\omega_\ell \mu_0} A_{\ell\nu}^* A_{\ell\mu} e^{j(\beta_{\ell\mu} - \beta_{\ell\nu})z} \mathbf{F}_{\ell\nu}^\dagger \mathbf{F}_{\ell\mu} \, \mathrm{d}x\mathrm{d}y. \tag{3.19}$$

The second summation vanishes due to eq. (3.15), while the terms inside the first summation are reduced to $|A_{\ell\mu}|^2$, meaning that the total power at frequency ω_ℓ is given by

$$P_{\mathrm{tot}}(\omega_\ell) = \sum_{\mu=1}^{M_\ell} |A_{\ell\mu}|^2 , \tag{3.20}$$

and that the magnitude square of the complex amplitude gives the power in Watts of each mode.

Using eq. (3.15), eq. (3.16) becomes

$$\sum_{\mu=1}^{M_\ell} \frac{\omega_\ell^2}{c^2} A_{\ell\mu} e^{j\beta_{\ell\mu} z} \iint_{-\infty}^{+\infty} \mathbf{F}_{\ell\nu}^\dagger \Delta\boldsymbol{\epsilon}_\ell \mathbf{F}_{\ell\mu} \, \mathrm{d}x\mathrm{d}y + 4j\omega_\ell \mu_0 \frac{\partial A_{\ell\nu}}{\partial z} e^{j\beta_{\ell\nu} z}. \tag{3.21}$$

Dividing by $4j\omega_\ell\mu_0$, introducing the following change of variables,

$$a_{\ell\mu} = A_{\ell\mu} e^{j\beta_{\ell\mu} z} , \tag{3.22a}$$

$$\frac{\partial A_{\ell\mu}}{\partial z} = e^{-j\beta_{\ell\mu}z}\left(\frac{\partial a_{\ell\mu}}{\partial z} - j\beta_{\ell\mu}a_{\ell\mu}\right), \tag{3.22b}$$

and rearranging the terms, we can highlight the z-derivative of the mode amplitude, the propagation constants, and perturbation tensor, obtaining

$$\frac{\partial a_{\ell\nu}}{\partial z} - j\beta_{\ell\nu}a_{\ell\nu} - j\sum_{\mu=1}^{M_\ell} a_{\ell\mu}\frac{\omega_\ell}{4}\iint_{-\infty}^{+\infty} \mathbf{F}_{\ell\nu}^{\dagger}\epsilon_0\Delta\boldsymbol{\epsilon}_\ell\mathbf{F}_{\ell\mu}\,\mathrm{d}x\mathrm{d}y\,. \tag{3.23}$$

We now split the permittivity perturbation in its real and imaginary parts:

$$\Delta\boldsymbol{\epsilon}_\ell = \Delta\tilde{\boldsymbol{\epsilon}}_\ell - j\boldsymbol{\epsilon}_{\boldsymbol{\ell}}''. \tag{3.24}$$

The real part of the permittivity perturbation inside the summation in eq. (3.23) now defines the well known linear coupling coefficients as the overlap integral [15, 82, 109, 117]

$$K_{\nu\mu}(\omega_\ell) = \epsilon_0\frac{\omega_\ell}{4}\iint_{-\infty}^{+\infty} \mathbf{F}_{\ell\nu}^{\dagger}\Delta\tilde{\boldsymbol{\epsilon}}_\ell\mathbf{F}_{\ell\mu}\,\mathrm{d}x\mathrm{d}y\,. \tag{3.25}$$

The imaginary part instead defines the fiber losses. Assuming the absence of MDL, and making use of eq. (3.15), the corresponding summation in eq. (3.23) is reduced to a single term reading

$$\frac{\omega_\ell}{4}\frac{2\omega_\ell\mu_0}{\beta_{\ell\nu}}\epsilon_0\boldsymbol{\epsilon}_\ell''a_{\ell\nu} \triangleq \frac{\alpha_\nu}{2}a_{\ell\nu}, \tag{3.26}$$

where α_ℓ is the power attenuation coefficient at frequency ℓ.

The linear propagation effects are finally completely described by the following equation, which constitutes the left-hand side of the side of the final equation

$$\frac{\partial a_{\ell\nu}}{\partial z} + \frac{\alpha_\ell}{2}a_\ell - j\beta_{\ell\nu}a_{\ell\nu} - j\sum_{\mu=1}^{M_\ell} K_{\nu\mu}(\omega_\ell)a_{\ell\mu}\,, \tag{3.27}$$

Equation (3.27) can also be written in matrix form, describing the evolution of the vector $\mathbf{a}_\ell$ of complex mode amplitudes as

$$\frac{\partial \mathbf{a}_\ell}{\partial z} + \frac{\alpha_\ell}{2}\mathbf{a}_\ell - j(\mathbf{B}_\ell + \mathbf{K}_\ell)\mathbf{a}_\ell\,, \tag{3.28}$$

where $\mathbf{B}_\ell$ is the diagonal matrix of propagation constants, and the elements of $\mathbf{K}_\ell$ are determined by eq. (3.25), finding the results of [117] with the addition of fiber losses.

3.2.2 *Nonlinear terms*

Here we complete the derivation of the propagation equations by carrying out the calculation of the third-order nonlinear polarization. Equation (3.4) can be written in a more compact form as

$$\mathbf{P}^{(3)} = p_1\mathbf{N}_1 + p_2\mathbf{N}_2 + p_3\mathbf{N}_3 + p_4\mathbf{N}_4 + p_5\mathbf{N}_5\,. \tag{3.29}$$

Remembering to divide by $j4\omega_\ell\mu_0$ as done for the linear part, the right-hand side of eq. (3.5) reads

$$\frac{j\omega_\ell}{4}\sum_{i=1}^{5} p_i\mathbf{N}_i \tag{3.30}$$

Substituting the electric field of eq. (3.6) in eq. (3.4), we obtain the expressions for each nonlinear term.

We begin by computing the dot product of eq. (3.4a), obtaining

$$\mathbf{E}_\ell \cdot \mathbf{E}_\ell = \left(\sum_{\mu=1}^{M_\ell} \mathbf{F}_{\ell\mu} A_{\ell\mu} e^{j\beta_{\ell\mu}z}\right) \cdot \left(\sum_{\eta=1}^{M_\ell} \mathbf{F}_{\ell\eta} A_{\ell\eta} e^{j\beta_{\ell\eta}z}\right) \tag{3.31}$$

$$= \sum_{\mu=1}^{M_\ell}\sum_{\eta=1}^{M_\ell} (\mathbf{F}_{\ell\mu}\cdot\mathbf{F}_{\ell\eta}) A_{\ell\mu}A_{\ell\eta} e^{j(\beta_{\ell\mu}+\beta_{\ell\eta})z} \tag{3.32}$$

Multiplying by $\mathbf{E}_\ell^*$ we can then write

$$(\mathbf{E}_\ell \cdot \mathbf{E}_\ell)\mathbf{E}_\ell^* = \left(\sum_{\mu=1}^{M_\ell}\sum_{\eta=1}^{M_\ell} (\mathbf{F}_{\ell\mu}\cdot\mathbf{F}_{\ell\eta}) A_{\ell\mu}A_{\ell\eta} e^{j(\beta_{\ell\mu}+\beta_{\ell\eta})z}\right) \cdot \sum_{\rho=1}^{M_\ell} \mathbf{F}_{\ell\rho}^* A_{\ell\rho}^* e^{-j\beta_{\ell\rho}z} \tag{3.33}$$

$$= \sum_{\mu=1}^{M_\ell}\sum_{\eta=1}^{M_\ell}\sum_{\rho=1}^{M_\ell} (\mathbf{F}_{\ell\mu}\cdot\mathbf{F}_{\ell\eta})\mathbf{F}_{\ell\rho}^* A_{\ell\mu}A_{\ell\eta}A_{\ell\rho}^* e^{j(\beta_{\ell\mu}+\beta_{\ell\eta}-\beta_{\ell\rho})z} \tag{3.34}$$

$$= \sum_{\mu=1}^{M_\ell}\sum_{\eta=1}^{M_\ell}\sum_{\rho=1}^{M_\ell} (\mathbf{F}_{\ell\mu}\cdot\mathbf{F}_{\ell\eta})\mathbf{F}_{\ell\rho}^* a_{\ell\mu}a_{\ell\eta}a_{\ell\rho}^*\,, \tag{3.35}$$

where in the last line we applied the change of variable of eq. (3.22).

Multiplying by $\mathbf{F}_{\ell\nu}^*$ as done for the left-hand side of the equation, and integrating on the transverse plane, we obtain the following

$$\sum_{\mu=1}^{M_\ell}\sum_{\eta=1}^{M_\ell}\sum_{\rho=1}^{M_\ell} a_{\ell\mu}a_{\ell\eta}a_{\ell\rho}^* \iint_{-\infty}^{+\infty} (\mathbf{F}_{\ell\mu}\cdot\mathbf{F}_{\ell\eta})(\mathbf{F}_{\ell\rho}^*\cdot\mathbf{F}_{\ell\nu}^*)\,\mathrm{d}x\mathrm{d}y\,. \tag{3.36}$$

Rewriting the overlap integral as

$$\mathcal{Q}^{(1)}_{\nu\rho\mu\eta}(\omega_\ell) = \iint_{-\infty}^{+\infty} (\mathbf{F}^*_{\ell\nu} \cdot \mathbf{F}^*_{\ell\rho})(\mathbf{F}_{\ell\mu} \cdot \mathbf{F}_{\ell\eta})\, \mathrm{d}x\mathrm{d}y\,, \tag{3.37}$$

we then obtain the expression of the first nonlinear term for mode ν, which can be written as

$$N_{1,\nu}(\omega_\ell) = \sum_{\mu=1}^{M_\ell}\sum_{\eta=1}^{M_\ell}\sum_{\rho=1}^{M_\ell} \mathcal{Q}^{(1)}_{\nu\rho\mu\eta} a_{\ell\mu} a_{\ell\eta} a^*_{\ell\rho} \tag{3.38}$$

The other nonlinear terms are similar and the expression for their overlap integrals are summarized in table 3.1. If the mode function of the modes with indices ρ, μ, or η is complex-conjugated inside the integral, then it must also be complex-conjugated in the triple summation when computing the total nonlinear contribution.

Regrouping the linear part of the equation with the nonlinear terms just described, the complete system of coupled nonlinear equations is finally obtained

$$\begin{aligned}
\frac{\partial a_{\ell\nu}}{\partial z} = &-\frac{\alpha_\ell}{2} a_{\ell\nu} + j\beta_{\ell\nu} a_{\ell\nu} + j\sum_{\mu=1}^{M_\ell} K_{\nu\mu}(\omega_\ell) a_{\ell\mu} \\
&+ \frac{j\omega_\ell}{4}\frac{\epsilon_0}{8}[\sigma + 2\tilde{b}_0] \sum_{\mu=1}^{M_\ell}\sum_{\eta=1}^{M_\ell}\sum_{\rho=1}^{M_\ell} \mathcal{Q}^{(1)}_{\nu\rho\mu\eta} a_{\ell\mu} a_{\ell\eta} a^*_{\ell\rho} \\
&+ \frac{j\omega_\ell}{4}\frac{\epsilon_0}{4}[\sigma + 2\tilde{a}_0 + \tilde{b}_0] \sum_{\mu=1}^{M_\ell}\sum_{\eta=1}^{M_\ell}\sum_{\rho=1}^{M_\ell} \mathcal{Q}^{(2)}_{\nu\rho\mu\eta}(\omega_\ell) a^*_{\ell\mu} a_{\ell\eta} a_{\ell\rho} \\
&+ \frac{j\omega_\ell}{4}\frac{\epsilon_0}{4}[\sigma + 2\tilde{a}_0 + \tilde{b}_\Omega] \sum_{\mu=1}^{M_f}\sum_{\eta=1}^{M_f}\sum_{\rho=1}^{M_\ell} \mathcal{Q}^{(3)}_{\nu\rho\mu\eta}(\omega_\ell) a^*_{f\mu} a_{f\eta} a_{\ell\rho} \\
&+ \frac{j\omega_\ell}{4}\frac{\epsilon_0}{4}[\sigma + \tilde{b}_0 + \tilde{b}_\Omega] \sum_{\mu=1}^{M_f}\sum_{\eta=1}^{M_\ell}\sum_{\rho=1}^{M_f} \mathcal{Q}^{(4)}_{\nu\rho\mu\eta}(\omega_\ell) a_{f\mu} a_{\ell\eta} a^*_{f\rho} \\
&+ \frac{j\omega_\ell}{4}\frac{\epsilon_0}{4}[\sigma + 2\tilde{a}_\Omega + \tilde{b}_0] \sum_{\mu=1}^{M_f}\sum_{\eta=1}^{M_\ell}\sum_{\rho=1}^{M_f} \mathcal{Q}^{(5)}_{\nu\rho\mu\eta}(\omega_\ell) a^*_{f\mu} a_{\ell\eta} a_{f\rho},
\end{aligned} \tag{3.39}$$

where we used the more compact subscript notation to indicate the frequency at which the parallel- and cross-polarized Raman responses $\tilde{a}$ and $\tilde{b}$ are evaluated.

3.2.3 *Nonlinear polarization parameters*

Ideally, we would prefer to describe the intensity of the nonlinear effects through the nonlinear refractive index parameter n_2, which is experimentally determined using a probe signal with a single frequency and polarization, rather rather than using σ. Assuming a signal with frequency ω_ℓ polarized along the $\mathbf{x}$ axis, and substituting it in eq. (3.4), only the contributions eqs. (3.4a) and (3.4b) become non-zero. The resulting polarization then reads:

$$\begin{aligned}\mathbf{P}_\ell^{(3)}(\omega_\ell) &= \frac{\epsilon_0}{8}\left(\sigma + 2\tilde{b}(0)\right) E_x^2 E_x^* \mathbf{x} + \frac{\epsilon_0}{4}\left(\sigma + 2a(\tilde{0}) + b(\tilde{0})\right) E_x^2 E_x^* \mathbf{x} \\ &= \frac{\epsilon_0}{4}\left(\frac{3\sigma}{2} + 2\tilde{a}(0) + 2\tilde{b}(0)\right) |E_x|^2 E_x \mathbf{x}\end{aligned} \tag{3.40}$$

Inserting this expression in eq. (3.5), and recalling that $\epsilon_\ell = n_\ell^2$, we obtain

$$\nabla^2 E_x \mathbf{x} + \frac{\omega_\ell}{c^2}\left[\epsilon_\ell + \frac{|E_x|^2}{4}\left(\frac{3\sigma}{2} + 2\tilde{a}(0) + 2\tilde{b}(0)\right)\right] E_x \mathbf{x} = 0\,. \tag{3.41}$$

Noting that σ, $\tilde{a}(0)$, and $\tilde{b}(0)$ are real, the term inside square brackets represents the total (linear and nonlinear) refractive index

$$n_{\mathrm{tot},\ell} = \sqrt{\epsilon_\ell + \frac{|E_x|^2}{4}\left(\frac{3\sigma}{2} + 2\tilde{a}(0) + 2\tilde{b}(0)\right)}. \tag{3.42}$$

Assuming the second term is small compared to the first, we can write its Taylor series expansion as

Table 3.1: Expression of the overlap integrals that define the strength of the nonlinear interaction between modes.

$$\begin{aligned}
\mathcal{Q}^{(1)}_{\nu\rho\mu\eta}(\omega_\ell) &= \int_{-\infty}^{+\infty}\!\!\int (\mathbf{F}^*_{\ell\nu} \cdot \mathbf{F}^*_{\ell\rho})(\mathbf{F}_{\ell\mu} \cdot \mathbf{F}_{\ell\eta})\,\mathrm{d}x\mathrm{d}y \\
\mathcal{Q}^{(2)}_{\nu\rho\mu\eta}(\omega_\ell) &= \int_{-\infty}^{+\infty}\!\!\int (\mathbf{F}^*_{\ell\nu} \cdot \mathbf{F}_{\ell\rho})(\mathbf{F}^*_{\ell\mu} \cdot \mathbf{F}_{\ell\eta})\,\mathrm{d}x\mathrm{d}y \\
\mathcal{Q}^{(3)}_{\nu\rho\mu\eta}(\omega_\ell) &= \int_{-\infty}^{+\infty}\!\!\int (\mathbf{F}^*_{\ell\nu} \cdot \mathbf{F}_{\ell\rho})(\mathbf{F}^*_{f\mu} \cdot \mathbf{F}_{f\eta})\,\mathrm{d}x\mathrm{d}y \\
\mathcal{Q}^{(4)}_{\nu\rho\mu\eta}(\omega_\ell) &= \int_{-\infty}^{+\infty}\!\!\int (\mathbf{F}^*_{\ell\nu} \cdot \mathbf{F}^*_{f\rho})(\mathbf{F}_{f\mu} \cdot \mathbf{F}_{\ell\eta})\,\mathrm{d}x\mathrm{d}y \\
\mathcal{Q}^{(5)}_{\nu\rho\mu\eta}(\omega_\ell) &= \int_{-\infty}^{+\infty}\!\!\int (\mathbf{F}^*_{\ell\nu} \cdot \mathbf{F}_{f\rho})(\mathbf{F}^*_{f\mu} \cdot \mathbf{F}_{\ell\eta})\,\mathrm{d}x\mathrm{d}y
\end{aligned}$$

TABLE I. Nonlinear polarization parameters defined in text as derived from our Raman data and intensity-induced polarization change (IIPC) measurements. From these are derived the self-focusing index n_2 of Eq. (9) and the fraction f_e of it that is electronic in origin. All quantities are in esu.

Material	n	$10^{16}B_0$ [a]	$10^{16}(A_0+B_0)$ [a]	$10^{15}(\sigma+2B_0)$ [b]	$10^{15}\sigma$	$10^{14}n_2$ [Eq. (4)]	$10^{14}n_2$ (Meas.)	f_e (%)
Fused quartz	1.46	14 ± 2	57 ± 6	31 ± 2	28 ± 2	11 ± 1	18 ± 3 [c]	79 ± 3
LSO glass	1.52	28 ± 3	70 ± 7	46 ± 3	40 ± 3	15 ± 1		81 ± 3
ED-4 glass	1.56	26 ± 3	68 ± 7	57 ± 4	52 ± 4	18 ± 1.3	26 ± 3 [c] 15 ± 2 [d] 15 ± 3 [e]	85 ± 3
SF-7 glass	1.65	210 ± 20	315 ± 30	202 ± 20	160 ± 20	58 ± 7		79 ± 3
LASF-7 glass	1.93	380 ± 40	540 ± 50	252 ± 25	176 ± 25	61 ± 8		71 ± 3

[a] This work. [b] References 1 and 12. [c] Reference 9. [d] Reference 8. [e] Reference 7.

Figure 3.1: Nonlinear polarization parameters reported in [118]

$$n_{\mathrm{tot},\ell} \simeq \sqrt{\epsilon_\ell} + \frac{1}{2\sqrt{\epsilon_\ell}} \frac{|E_x|^2}{4} \left(\frac{3\sigma}{2} + 2\tilde{a}(0) + 2\tilde{b}(0) \right) \tag{3.43}$$

$$= n_\ell + \frac{|E_x|^2}{8n_\ell} \left(\frac{3\sigma}{2} + 2\tilde{a}(0) + 2\tilde{b}(0) \right) = n_\ell + n_{2,\ell}|E_x|^2, \tag{3.44}$$

where the nonlinear refractive index is defined as

$$n_{2,\ell} = \frac{1}{8n_\ell} \left(\frac{3\sigma}{2} + 2\tilde{a}(0) + 2\tilde{b}(0) \right) \tag{3.45}$$

This quantity, however, is expressed in $\mathrm{m^2\,V^{-2}}$, while it is more customary to defined it in units of $\mathrm{m^2\,W^{-1}}$. In order to convert one to the other, we can exploit its relationship with the average intensity

$$I = \frac{|E_x|^2}{2\eta}, \tag{3.46}$$

with

$$\eta = \sqrt{\frac{\mu_0}{\epsilon\epsilon_0}} = \sqrt{\frac{\mu_0\epsilon_0}{\epsilon\epsilon_0^2}} = \frac{1}{c\epsilon_0 n}. \tag{3.47}$$

Using the relation

$$n_2|E_x|^2 = n_{2,\mathrm{W}} \frac{|E_x|^2}{2\eta}, \tag{3.48}$$

it follows that the nonlinear refractive index expressed in units of $\mathrm{m^2\,W^{-1}}$ is defined as

$$n_{2,\mathrm{W}} = \frac{2n_2}{c\epsilon_0 n} = \frac{3\sigma/2 + 2\tilde{a}(0) + 2\tilde{b}(0)}{4c\epsilon_0 n^2}. \tag{3.49}$$

Some reference values for the nonlinear polarization parameters are reported in Table I of [118], which is reported in fig. 3.1.

3.2.3.1 *Raman gain coefficient*

Let us consider the electric fields for pump and signal linearly polarized along the x-axis:

$$\mathbf{E}_p(t) = \mathrm{Re}\Big\{E_{p,x}e^{-j\omega_p t}\mathbf{x}\Big\} \tag{3.50}$$

$$\mathbf{E}_s(t) = \mathrm{Re}\Big\{E_{s,x}e^{-j\omega_s t}\mathbf{x}\Big\} \tag{3.51}$$

The nonlinear polarization at the signal frequency is found by summing the terms of eqs. (3.4c) to (3.4e) and reads

$$\mathbf{P}_s^{(3)} = \frac{\epsilon_0}{4}\left(3\sigma + 2\tilde{a}(0) + 2\tilde{b}(0) + 2\tilde{a}(\omega_p-\omega_s) + 2\tilde{b}(\omega_p-\omega_s)\right)|E_{p,x}|^2E_{s,x}\mathbf{x}\,. \tag{3.52}$$

The imaginary part of the previous equation determines the Raman gain, meaning that σ, $\tilde{a}(0)$, and $\tilde{b}(0)$ disappear, and the total refractive index defined in eq. (3.43) becomes complex

$$n_C \simeq n + n_2|E_{s,x}|^2 + j\frac{1}{4}\,\mathrm{Im}\big\{2\tilde{a}(\omega_p-\omega_s) + 2\tilde{b}(\omega_p-\omega_s)\big\}|E_{p,x}|^2\,. \tag{3.53}$$

The propagation factor for a wave propagating in the z-direction is written as

$$\exp\Big(-j\frac{\omega_s n_C}{c}z\Big) \simeq \exp\Big[-j\omega_s(n+n_2|E_{sx}|^2)\frac{z}{c}\Big] \tag{3.54}$$

$$\exp\left[-j\frac{\omega_s}{c}\frac{j}{4}\,\mathrm{Im}(\tilde{a}(\omega_p-\omega_s)+\tilde{b}(\omega_p-\omega_s))|E_{p,x}|^2z\right] \tag{3.55}$$

$$= \exp(-jk_{NL}z)\exp(gz/2)\,, \tag{3.56}$$

where g is the gain coefficient for the electric field amplitude in units of m^{-1}, defined as

$$g = \frac{\omega_s}{2c}\,\mathrm{Im}[\tilde{a}(\omega_p-\omega_s)+\tilde{b}(\omega_p-\omega_s)]|E_{p,x}|^2\,. \tag{3.57}$$

We need to relate this coefficient to the familiar Raman gain coefficient g_R in units of $\mathrm{m\,W}^{-1}$, using the definition of η in eq. (3.47)

$$g = \eta\frac{\omega_s}{c}\,\mathrm{Im}[\tilde{a}(\omega_p-\omega_s)+\tilde{b}(\omega_p-\omega_s)]\frac{|E_{p,x}|^2}{2\eta}\,. \tag{3.58}$$

and obtain the expression for g_R as

$$g_R = \eta\frac{\omega_s}{c}\,\mathrm{Im}[\tilde{a}(\omega_p-\omega_s)+\tilde{b}(\omega_p-\omega_s)] \tag{3.59}$$

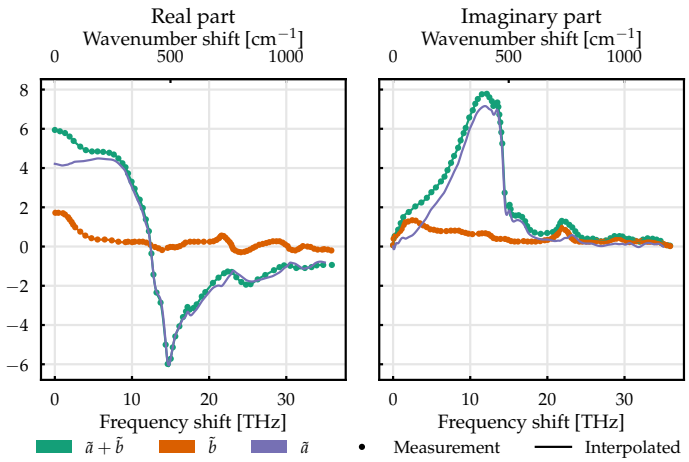


Figure 3.2: Measured Raman susceptibility curves and their interpolation. Data from [119].

.

$$= \frac{\omega_s}{c^2 \epsilon_0 n} \operatorname{Im}[\tilde{a}(\omega_p - \omega_s) + \tilde{b}(\omega_p - \omega_s)] \tag{3.60}$$

The shape of the Raman response functions $\tilde{a}$ and $\tilde{b}$ can be directly obtained by experimental measurements, from e.g. [119]. Their shape is illustrated in fig. 3.2. For simulation purposes, these curves are appropriately scaled to obtain a desired Raman gain coefficient g_R. They also determine the value of the parameter σ that is required to get a target nonlinear refractive index through eq. (3.49),

3.2.4 *Single-mode case*

In order to assess the correctness of the model it is useful to compare it to the well known equations for forward-pumped Raman amplification in a single mode fiber, which read [20]

$$\frac{dP_s}{dz} = \frac{g_R}{A_{\text{eff}}} P_s P_p - \alpha_s P_s \tag{3.61a}$$

$$\frac{dP_p}{dz} = -\frac{\omega_p}{\omega_s} \frac{g_R}{A_{\text{eff}}} P_s P_p - \alpha_p P_p \,, \tag{3.61b}$$

where P_p and P_s represent the power at the pump and signal frequency, respectively, g_R is the Ramain gain coefficient as defined in eq. (3.59), and A_{eff} is the effective core area of the fiber. In the undepleted pump approximation, the solution for the signal power is obtained analytically as

$$P_s(z) = P_s(0) \exp\left[\frac{g_R}{A_{\text{eff}}} P_p(0) \frac{1 - e^{-\alpha_p z}}{\alpha_p} - \alpha_s z\right] . \tag{3.62}$$

Adapting eq. (3.39) for the case of single-mode co-polarized signal and pump waves, neglecting linear mode coupling and the first two nonlinear terms which only account for instantaneous Kerr nonlinearity, we can write the following equations, considering the undepleted pump case

$$\frac{\partial a_s}{\partial z} = -\frac{\alpha_s}{2}a_s + j\beta_s a_s \tag{3.63}$$

$$+ \frac{j\omega_\ell}{4}\frac{\epsilon_0}{4}[\sigma + 2\tilde{a}_0 + \tilde{b}_\Omega]\mathcal{Q}^{(3)}_{1111}(\omega_s)a_p^* a_p a_s \tag{3.64}$$

$$+ \frac{j\omega_\ell}{4}\frac{\epsilon_0}{4}[\sigma + \tilde{b}_0 + \tilde{b}_\Omega]\mathcal{Q}^{(4)}_{1111}(\omega_s)a_p a_s a_p^* \tag{3.65}$$

$$+ \frac{j\omega_\ell}{4}\frac{\epsilon_0}{4}[\sigma + 2\tilde{a}_\Omega + \tilde{b}_0]\mathcal{Q}^{(5)}_{1111}(\omega_s)a_p^* a_s a_p \tag{3.66}$$

$$\frac{\partial a_p}{\partial z} = -\frac{\alpha_p}{2}a_p + j\beta_p a_p \,. \tag{3.67}$$

Supposing that the mode field functions do not vary much in frequency, we can consider $\mathbf{F}(\omega_s) \approx \mathbf{F}(\omega_p)$, and consequently $\mathcal{Q}^{(3)}_{1111}(\omega_s) = \mathcal{Q}^{(4)}_{1111}(\omega_s) = \mathcal{Q}^{(5)}_{1111}(\omega_s) = \mathcal{Q}_{1111}$. The previous equations then become

$$\frac{\partial a_s}{\partial z} = -\frac{\alpha_s}{2}a_s + j\beta_s a_s + \frac{j\omega_s}{8}\epsilon_0\left[\frac{3\sigma}{2} + \tilde{a}_0 + \tilde{b}_0 + \tilde{a}_\Omega + \tilde{b}_\Omega\right]\mathcal{Q}_{1111}|a_p|^2 a_s \tag{3.68}$$

$$\frac{\partial a_p}{\partial z} = -\frac{\alpha_p}{2}a_p + j\beta_p a_p \,. \tag{3.69}$$

The solution for the pump equation is straightforward, and leads to the following expression for the evolution of the pump power

$$P_s(z) = |a_s(z)|^2 = |a_{p0}|^2 \exp\left(-\alpha_p z\right) \,. \tag{3.70}$$

Substituting eq. (3.70) in the equation for the signal instead we obtain

$$\frac{\partial a_s}{\partial z} = -\frac{\alpha_s}{2}a_s + j\beta_s a_s + G|a_{p0}|^2 \exp\left(-\alpha_p z\right) a_s \,, \tag{3.71}$$

with $G = j\omega_s/8\epsilon_0\left[3\sigma/2 + \tilde{a}_0 + \tilde{b}_0 + \tilde{a}_\Omega + \tilde{b}_\Omega\right]\mathcal{Q}_{1111}$, and which leads to the following solution for the complex amplitude

$$a_s(z) = \exp\left(-\frac{\alpha_s}{2}z + j\beta_s z\right)\int_0^{z'} G|a_{p0}|^2 \exp\left(-\alpha_p z'\right)\mathrm{d}z' \tag{3.72}$$

$$= \exp\left(-\frac{\alpha_s}{2}z + j\beta_s z\right)\exp\left(G|a_{p0}|^2 \cdot \frac{1 - e^{-\alpha_p z}}{\alpha_p}\right)a_{s0}. \tag{3.73}$$

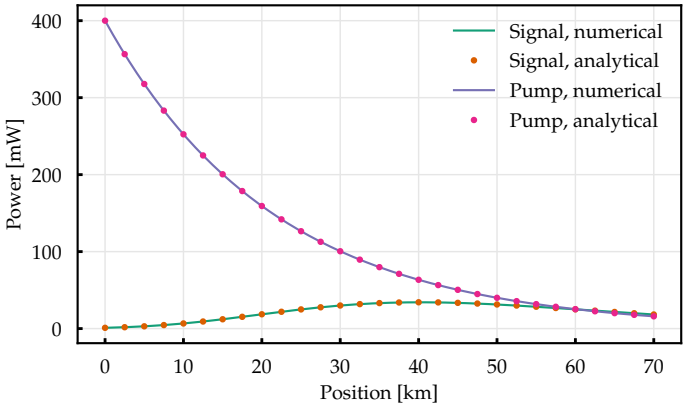


Figure 3.3: Comparison between analytical expression and numerical solution.

The signal power becomes

$$P_s(z) = |a_s(z)|^2 = a_s a_s^* = |a_{s0}|^2 \exp\left(2\,\mathrm{Re}[G]|a_{p0}|^2 \cdot \frac{1-e^{-\alpha_p z}}{\alpha_p} - \alpha_s z\right), \tag{3.74}$$

and considering that σ, $\tilde{a}_0$, $\tilde{b}_0$, and $\mathcal{Q}_{1111}$ (for the case of LP modes) are real numbers, and $\mathrm{Re}[G] = -\omega_s/8\epsilon_0\mathcal{Q}_{1111}\,\mathrm{Im}[\tilde{a}_\Omega + \tilde{b}_\Omega]$, the final expression for the signal power becomes

$$P_s(z) = |a_{s0}|^2 \exp\left(-\frac{\omega_s}{4}\epsilon_0\mathcal{Q}_{1111}\,\mathrm{Im}[\tilde{a}_\Omega + \tilde{b}_\Omega]|a_{p0}|^2 \cdot \frac{1-e^{-\alpha_p z}}{\alpha_p} - \alpha_s z\right). \tag{3.75}$$

Comparing with eq. (3.62), it is easy to see that

$$\frac{g_R}{A_{\mathrm{eff}}} = -\frac{\omega_s}{4}\epsilon_0\mathcal{Q}_{1111}\,\mathrm{Im}[\tilde{a}_\Omega + \tilde{b}_\Omega]. \tag{3.76}$$

In fig. 3.3 we compare the analytical solution given in eq. (3.75) with the results of numerical integration of the derived model without considering linear coupling, launching power in the x-polarization of the LP_{01} mode of a 70 km long fiber supporting two mode groups, using a peak Raman gain coefficient of $g_R = 1 \times 10^{-13}\,\mathrm{m\,W^{-1}}$.

3.3 NUMERICAL METHODS

In order to study the propagation of the different modes of a FMF under the effect of intrinsic and extrinsic perturbations, a suitable numerical model must be defined. In this section we describe how the perturbations acting on a ideal fiber are modeled, and outline

the algorithm used to solve the propagation equations derived in the previous section for the resulting forward-pumping Raman amplifier.

3.3.1 *Linear propagation*

The linear effects occurring during propagation in a FMF are accounted for by the linear terms of eq. (3.39), describing the evolution of the vector of the field complex amplititudes, and can be written in matrix form as

$$\frac{\partial \mathbf{a}}{\partial z} = -\frac{\alpha}{2}\mathbf{a} + j(\mathbf{B} + \mathbf{K})\mathbf{a}\,, \tag{3.77}$$

which is the same used in [15, 82] with the addition of fiber losses.

For simplicity, we momentarily drop the subscript notation to indicate the frequency dependence of each quantity, as the linear effects, at least in first approximation, can be treated individually for each frequency component.

The local coupling effects, caused by perturbations to the ideal fiber profile, are described by the coupling matrix $\mathbf{K}$, while $\mathbf{B}$ is the diagonal matrix of propagation constants. The coupling strength of a specific perturbation can be calculated from the eigenvalues κ_i of its coupling matrix and related to the its coupling beat length L_κ as

$$L_\kappa = \frac{2\pi}{\Delta\kappa}, \tag{3.78}$$

where $\Delta\kappa = \max_i \kappa_i - \min_i \kappa_i$ is the coupling strength.

In the most general case, both the strength and the orientation of the various perturbations can vary randomly along the fiber. It is then necessary to account for this by explicitly highlighting the dependence of the coupling matrix on the position along the fiber. If we consider the normalized coupling matrices $\bar{\mathbf{K}}$, computed with the reference perturbation aligned to the reference frame of the fiber, and scaled such that their strength is equal to 1, the total coupling matrix of the fiber, with multiple perturbations acting on it, reads [82]

$$\mathbf{K}(z) = \sum_\zeta \Gamma_\zeta(z)\mathbf{R}(\theta_\zeta(z))\bar{\mathbf{K}}^{(\zeta)}\mathbf{R}^\top(\theta_\zeta(z)), \tag{3.79}$$

where Γ_ζ and $\theta_\zeta(z)$ are the strength and angle of perturbation ζ. Matrix $\mathbf{R}$ is a rotation operator, which, by definition, does not change the strength of the perturbation. The form of matrix $\mathbf{R}$ is that of a block diagonal matrix, where each block corresponds to a different manifold of degenerate modes [109]. Hereinafter, in each group of degenerate modes, we order the modes by alternating x and y polarizations of the even degeneracy, and then the x and y polarizations of odd degeneracy.

With this notation, for manifolds with azimuthal order 0, e.g. the LP01 group, the corresponding block can be written as [15]

$$\mathbf{R}_0(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{3.80}$$

while for manifolds of order n, it reads [15]

$$\mathbf{R}_n(\theta) = \begin{pmatrix} \mathbf{R}_0(\theta) & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_0(\theta) \end{pmatrix} \begin{pmatrix} \cos(n\theta)\mathbf{I} & -\sin(n\theta)\mathbf{I} \\ \sin(n\theta)\mathbf{I} & \cos(n\theta)\mathbf{I} \end{pmatrix}, \tag{3.81}$$

where $\mathbf{0}$ and $\mathbf{I}$ are the 2×2 zero-valued and identity matrices, respectively.

Here, as in [82], we focus only on the effect of stress birefringence and core ellipticity, which are perturbations that are closely related one another, as they are mainly due to the fiber manufacturing process. With this consideration, we assume that they act on the fiber with the same perturbation angle, and contribute equally to the overall strength [120, 121]. Following this hypothesis, the total coupling matrix for the case of stress birefringence and core ellipticity can be simplified and written as

$$\mathbf{K}(z) = \Gamma(z)\mathbf{R}(\theta(z))\bar{\mathbf{K}}^{(t)}\mathbf{R}^{\top}(\theta(z)), \tag{3.82}$$

where $\mathbf{K}^{(t)} = \bar{\mathbf{K}}^{(e)} + \bar{\mathbf{K}}^{(b)}$, is the total, unnormalized coupling matrix, while $\bar{\mathbf{K}}^{(e)}$ and $\bar{\mathbf{K}}^{(b)}$ are the normalized coupling matrices for core ellipticity and stress birefringence, respectively.

In order to model the evolution of these perturbations along the fiber, we use the well-known fixed-modulus model (FMM), which is commonly used in the study PMD in SMFs [122]. According to this model, the strength of the perturbation remains constant along the fiber, while their angle θ varies according to a Wiener process [122]

$$\frac{d\theta}{dz} = -\sigma_\theta w(z), \tag{3.83}$$

where w is a Gaussian white noise process with zero mean and unit variance, and σ_θ is related to the correlation length L_c of the process [122]

$$\sigma_\theta = \frac{1}{\sqrt{2L_c}}. \tag{3.84}$$

Recalling that the multiplication of a matrix by a scalar has the effect of scaling its eigenvalues by the same amount, we can set the strength of the total perturbation, or equivalently, the coupling beat length, by

simply multiplying the total normalized matrix $\bar{\mathbf{K}}^{(t)}$ by the appropriate amount. Under the FMM, this translates to setting

$$\Gamma(z) = \Gamma_0 = \frac{2\pi}{L_\kappa}. \tag{3.85}$$

The use of the FMM is justified when considering intrinsic perturbations such as birefringence and core ellipticity resulting from fiber manufacturing; it is in fact reasonable to assume that the drawing process is stable enough to induce negligible birefringence variations. On the other hand, for extrinsic perturbations like e.g. twisting or bending that depend on the deployment and installation of the fiber cable, models that account for the change of the perturbation strength along the fiber (such as the random modulus model) are more appropriate, and better capture the localized nature of some of these effects.

The most common way to numerically model the evolution of the perturbations along the fiber is to discretize the Wiener process defining the angle of the perturbations. This is accomplished by dividing the optical fiber in a series of N_s plates, each of size $\delta z = L/N_s$, where L is the total length of the fiber. [82, 112, 123–125]. The number of plates must be sufficiently high so that the angle $\theta(z)$ can be considered almost constant over their length, meaning that $\delta z \ll L_c$.

Over the length of the k-th plate, the propagation equation then assumes the following form

$$\frac{\partial \mathbf{a}}{\partial z} = \mathbf{L}(\theta_k)\mathbf{a} \qquad z \in [k\delta z, (k+1)\delta z], \tag{3.86}$$

with

$$\mathbf{L}(\theta_k) = -\frac{\alpha}{2} + j\left[\mathbf{B} + \frac{2\pi}{L_\kappa}\mathbf{R}(\theta_k))\bar{\mathbf{K}}^{(t)}\mathbf{R}^\top(\theta_k)\right]. \tag{3.87}$$

This means that the propagation over a single plate has a closed form solution written as

$$\mathbf{a}(k+1) = \exp[\mathbf{L}(\theta_k)\delta z]\mathbf{a}(k). \tag{3.88}$$

This fact brings a significant advantage when dealing with the numerical solution of the linear part of eq. (3.39): if this were not the case, we would need to apply conventional integration schemes such as the Euler or Runge-Kutta scheme with a step size that is substantially smaller than the minimum beat length involved, which is a least $L_\beta = 2\pi/\Delta\beta$, resulting in integration steps of fractions of millimeters.

Numerically, the matrix exponential in eq. (3.88) can be computed with a single eigenvalue decomposition. Indeed, recalling that $\mathbf{B}$ is diagonal, that the rotation matrices $\mathbf{R}$ are orthogonal, that $\mathbf{K}$ is Hermi-

tian and thus diagonalizable, and that for two $n \times n$ complex matrices $\mathbf{X}$, $\mathbf{Y}$ it holds that [126]

$$\exp\left[\mathbf{X}\mathbf{Y}\mathbf{X}^{-1}\right] = \mathbf{X}\exp[\mathbf{Y}]\mathbf{X}^{-1}, \tag{3.89}$$

we can solve eq. (3.88) by computing the eigenvalues λ_i and the matrix of eigenvectors $\mathbf{M}$ of $\mathbf{L}(\theta_k)$, obtaining

$$\exp[\mathbf{L}(\theta_k)\delta z] = \mathbf{M}\exp[\mathbf{\Lambda}\delta z]\mathbf{M}^{-1} \tag{3.90}$$

with $\mathbf{\Lambda} = \mathrm{diag}(\lambda_0, \ldots, \lambda_{n-1})$. As a consequence, the step size required to accurately follow the evolution of the vector of complex amplitudes should only be a fraction of the correlation length of the Wiener process describing the orientation of the perturbation, which is usually in the order of tens of meters, allowing us to significantly speed up the numerical integration of the propagation equations.

3.3.2 *Nonlinear propagation*

In the previous section we introduced the theoretical and numerical model for the propagation of multiple modes of a FMF in the presence of linear propagation effects, highlighting how a closed-form solution can be determined on sub-intervals of the integration domain. Apart from specific and simplified cases, this is not true when dealing with only nonlinear phenomena, more so when both linear and nonlinear effects act together during propagation.

The typical approach in such cases is to employ the so-called split-step integration schemes [114, 127–129]. In this framework, we assume that in the propagation over a short portion of the fiber of length h, linear and nonlinear effects act independently. As such, the propagation from position z to $z + h$ can be carried out in two separate steps. In the first step, linear effects act alone, and the propagation is solved using the methods detailed in the previous section. In the second step, the nonlinear operator is applied on the linearly propagated field computed in the first step.

Helped by the fact that the length scale of nonlinear effects is typically much longer than that of linear effects [103, 125, 130], we can safely use the same step size employed for the sole linear propagation, that is, fractions of the perturbation orientation correlation length, as done, for example, in [125].

Rewriting eq. (3.39) for the pump and signal frequencies ω_p and ω_s, and collecting its linear and nonlinear components, the following system of equations is found

$$\frac{\partial \mathbf{a}(z;\omega_s)}{\partial z} = \mathcal{L}\left[\mathbf{a}(z;\omega_s)\right] + \mathcal{N}\left[\mathbf{a}(z;\omega_s), \mathbf{a}(z;\omega_p)\right] \tag{3.91}$$

$$\frac{\partial \mathbf{a}(z;\omega_p)}{\partial z} = \mathcal{L}\left[\mathbf{a}(z;\omega_p)\right] + \mathcal{N}\left[\mathbf{a}(z;\omega_p), \mathbf{a}(z;\omega_s)\right], \tag{3.92}$$

where $\mathcal{L}$ and $\mathcal{N}$ contain all the linear and nonlinear contributions, respectively. Using the split-step integration algorithm, the propagation over the k-th plate of the fiber is then found to be

$$\mathbf{a}^{(L)}(k+1;\omega_s) = \exp[\mathbf{L}(\theta_k;\omega_s)\delta z]\mathbf{a}(k;\omega_s) \tag{3.93a}$$

$$\mathbf{a}^{(L)}(k+1;\omega_p) = \exp\left[\mathbf{L}(\theta_k;\omega_p)\delta z\right]\mathbf{a}(k;\omega_p) \tag{3.93b}$$

$$\mathbf{a}(k+1;\omega_s) = \exp\left[\mathcal{N}\left[\mathbf{a}^{(L)}(k+1;\omega_s), \mathbf{a}^{(L)}(k+1;\omega_p)\right]\delta z\right] \tag{3.93c}$$

$$\mathbf{a}(k+1;\omega_p) = \exp\left[\mathcal{N}\left[\mathbf{a}^{(L)}(k+1;\omega_p), \mathbf{a}^{(L)}(k+1;\omega_s)\right]\delta z\right] \tag{3.93d}$$

Numerically, eq. (3.93a) and eq. (3.93b) are solved using the formula in eq. (3.90), while the nonlinear operators in eq. (3.93c) and eq. (3.93d) are computed with a fourth-order Runge-Kutta scheme. It should be noted that in this formulation, co-propagation of signal and pump is assumed.

3.4 RESULTS

As detailed in the previous section, we consider a few-mode fiber Raman amplified link in which only two wavelengths are propagating; one acting as the Raman pump, and the other consisting of the information-bearing signal to be amplified, each in the constant-wave (CW) regime. The signal frequency corresponds to the wavelength $\lambda_s = 1550\,\text{nm}$, while the pump is frequency up shifted by 12 THz, approximately corresponding to the frequency detuning for which the imaginary part of the Raman response is maximized.

We simulate two different FMFs, each with a step-index profile, supporting 2 (LP_{01} and LP_{11}) and 4 (LP_{01}, LP_{11}, LP_{21}, and LP_{02}) groups of LP modes, respectively. Both fibers have a cladding diameter equal to $d_{cl} = 120\,\mu\text{m}$ and a core diameter equal to $d_{co} = 12\,\mu\text{m}$; the core refractive index set to $n_{co} = 1.46$, while their cladding refractive index n_{cl} is calculated to ensure that the number of supported modes at signal and pump frequencies is the same.

The two fiber geometries have been simulated using the commercial finite-element method (FEM) solver COMSOL ©Multiphysics [131], which computes the propagation constant and the field distribution of each guided mode.

As the true modes supported by the fiber belong to the family of hybrid modes, a standard combination procedure is applied to convert them to LP modes [98, 132]. Additionally, since the solver uses a reference frame which is rotated by an arbitrary angle for the computation of each mode, an additional orthogonalization step is applied to align the modes on a common Cartesian reference frame. Finally, the mode distribution functions are normalized according to eq. (3.15).

The integrals in eq. (3.25) and table 3.1 are then numerically evaluated for both signal and pump frequencies to obtain the linear and nonlinear coupling coefficients, respectively.

3.4.1 *Linear coupling matrices*

As previously introduced, we only consider the combined effect of stress birefringence and core ellipticity on the overall propagation dynamics.

Their effect is included by properly modeling the tensor of dielectric perturbation $\Delta\tilde{\epsilon}$ which appears in the expression of the linear coupling coefficients of eq. (3.25). Depending on the form of the tensor, coupling between different modes can occur either among transverse components of the field, resulting in a strong interaction, or between longitudinal components, causing weaker effects.

3.4.1.1 *Birefringence*

For the case of stress birefringence, the tensor of perturbation reads [15, 109]

$$\Delta\tilde{\epsilon}(x,y) = n(x,y)\delta n \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{3.94}$$

where $n(x,y)$ and δn are the refractive index distribution of the fiber and the fiber birefringence, respectively. The explicit computation of the coupling coefficients reveals that birefringence is simply responsible for the detuning of the x- and y- polarization of each spatial mode [15]. Consequently, due to orthogonality conditions, coupling only occurs between transverse components of the modes belonging to the

same group [109]. In fact, the theoretical coupling matrices for groups $\mathrm{LP}_{0,p}$ and $\mathrm{LP}_{n,p}$, with $n > 0$, read

$$\mathbf{K}_{0,p} = c_{0,p} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \mathbf{K}_{n,p} = c_{n,p} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{3.95}$$

where $c_{n,p}$ are properly defined coefficients [15].

3.4.1.2 *Core ellipticity*

In the case of an elliptical core, the perturbation in polar coordinates reads [15, 109]

$$\Delta\tilde{\epsilon}(r,\phi) \simeq \gamma(n_{\mathrm{co}}^2 - n_{\mathrm{cl}}^2)\delta(r - r_{\mathrm{co}})\cos 2\phi, \tag{3.96}$$

where γ represents the overall ellipticity in terms of maximum radius variation and $\delta(\cdot)$ is the Dirac function. As detailed in [109], core ellipticity causes coupling between the transverse components among modes with azimuthal order 1 and among modes whose azimuthal order differs by 2. Coupling between longitudinal components instead occurs among modes with the same azimuthal order or with orders differing by 4. This means that in a fiber supporting only the LP_{01} and LP_{11} groups, no inter-group coupling occurs. Differently, if the fiber also supports the propagation of the LP_{21} and LP_{02} groups, strong inter-group coupling is expected for the LP_{01}-LP_{21} and LP_{02}-LP_{21} pairs, while weak inter-group coupling is predicted between the LP_{01} and LP_{02} groups.

3.4.1.3 *Numerical computation*

From the theoretical formulation of the tensor of perturbation, the linear coupling integrals are numerically evaluated for both the pump and the signal frequency, using reference values of birefringence $\bar{\delta n}$ and ellipticity $\bar{\gamma}$.

Some post-processing steps are also applied:

1. given that the coupling coefficients must be real for both birefringence and core ellipticity, the residual imaginary part of the corresponding matrices is set to 0;
2. following eq. (3.25), Hermitian symmetry is imposed;
3. the coupling rules detailed in section 3.4.1.1 and section 3.4.1.2 are enforced in order to remove spurious interaction between uncoupled modes;

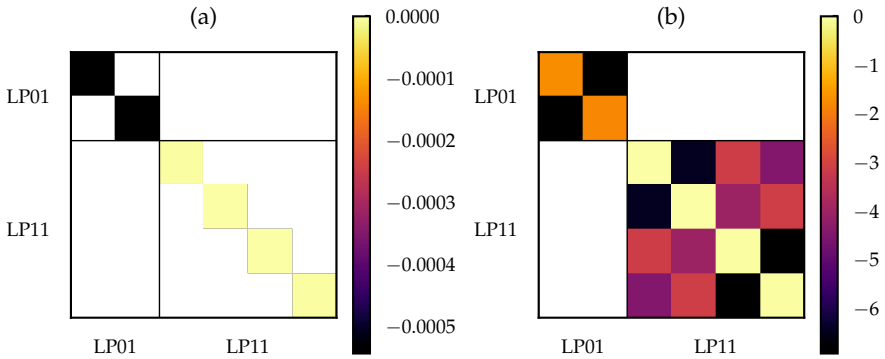


Figure 3.4: Linear coupling matrices for the 2-modes fiber for (a) stress birefringence and (b) core ellipticity. Values are reported in logarithmic scale.

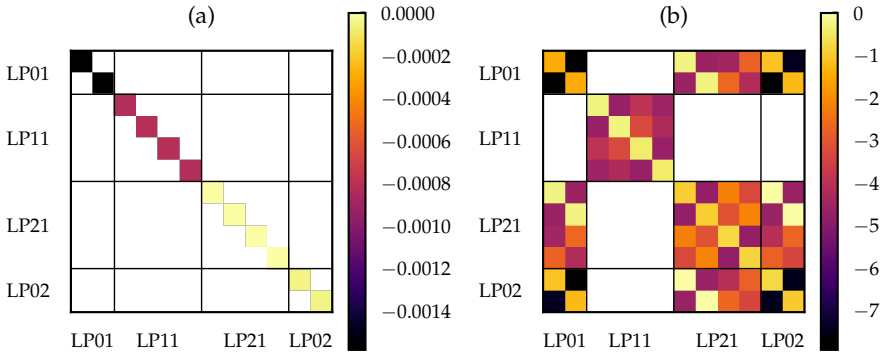


Figure 3.5: Linear coupling matrices for the 4-modes fiber for (a) stress birefringence and (b) core ellipticity. Values are reported in logarithmic scale.

4. in order to ensure that the computed matrices have unit strength, they are then scaled by the difference between their biggest and smallest eigenvalues, i.e. their coupling strength. As the perturbation strength depends linearly on δn and γ, the reference birefringence and ellipticity values can be scaled by the same amount.

The absolute value of the resulting linear coupling matrices for the 2-modes (4-modes) fiber are reported in logarithmic scale in fig. 3.4 (fig. 3.5).

Once the normalized matrices $\bar{\mathbf{K}}^{(e)}$ and $\bar{\mathbf{K}}^{(b)}$ are obtained, along with their reference birefringence $\bar{\delta n}$ and ellipticity $\bar{\gamma}$, the total normalized linear coupling matrix $\bar{\mathbf{K}}^{(t)}$ is obtained by adding the two matrices and normalizing it by its coupling strength.

This matrix is computed once for each frequency, and scaled accordingly during simulation to obtain the desired total coupling strength $\Delta\kappa$ or beat length L_κ.

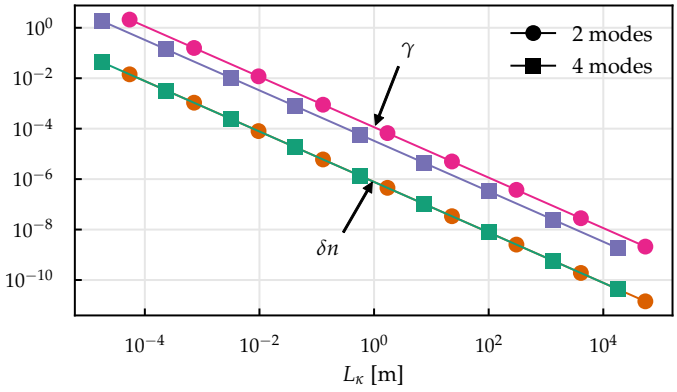


Figure 3.6: Equivalent birefringence δn and core ellipticity γ for a given coupling beat length L_κ, for the two fibers considered.

From this value, the corresponding birefringence and ellipticity parameters are retrieved using the approximation

$$\Delta\kappa \approx \Delta\kappa_e + \Delta\kappa_b \implies \Delta\kappa_e = \Delta\kappa_b \approx \frac{1}{2}\Delta\kappa, \tag{3.97}$$

where $\Delta\kappa = 2\pi/L_\kappa$ is the overall coupling strength, obtaining

$$\delta n \approx \frac{\pi}{L_\kappa}\bar{\delta n}, \qquad \gamma \approx \frac{\pi}{L_\kappa}\bar{\gamma}. \tag{3.98}$$

Using these relations, we can simply map a given coupling beat length to the corresponding physical parameters. They are illustrated in fig. 3.6 for a wide range of possible beat lengths.

The approximation is validated numerically, determining that it introduces relative error of ≈ 1 % on the total beat length.

3.4.2 *Optimization of the nonlinear step*

As pointed out in [99], depending on the symmetry properties of the particular fiber that is being studied, many of the overlap integrals $\mathcal{Q}^{(i)}_{\nu\rho\mu\eta}$ compute to 0, not contributing to the propagation dynamics.

This detail can help us optimize the integration algorithm presented earlier by reducing the number of combinations of modes to compute at each integration step, which is the dominating contribution to the total complexity of the numerical solver.

Since the mode distributions are computed with a FEM solver, numerical errors can break their symmetry, resulting in overlap integrals that are not identically equal to 0. We tackle the issue of identifying which coefficients should be discarded by following the approach

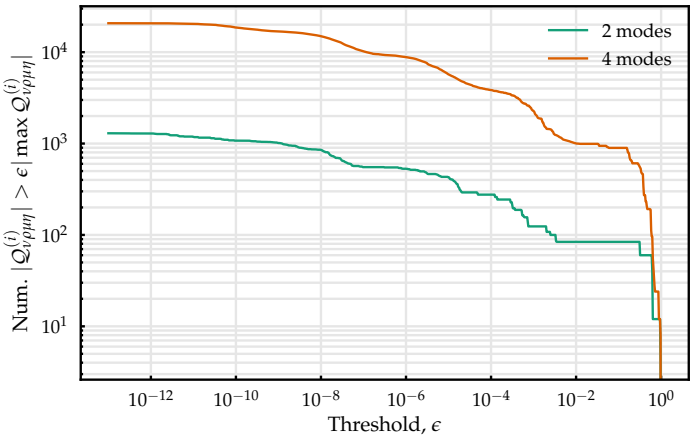


Figure 3.7: Statistics of the overlap integrals for the two considered fibers. Represented is the maximum with respect to the index i, at the signal frequency ω_s, corresponding to $\lambda_s = 1550$ nm.

.

detailed in [99], comparing the relative magnitude of the coefficients to their maximum value. Given a threshold coefficient ϵ, and the index i defining which of the 5 nonlinear terms we are considering, we only compute the nonlinear contributions for the overlap integrals that satisfy

$$|\mathcal{Q}^{(i)}_{\nu\rho\mu\eta}| > \epsilon \max_{\nu\rho\mu\eta} |\mathcal{Q}^{(i)}_{\nu\rho\mu\eta}|, \qquad i \in 1, \dots, 5. \tag{3.99}$$

As an example, by setting $\epsilon = 0.01$, we only consider the overlap integrals which are, in magnitude, greater than 1 percent of the biggest coefficient. This metric is visualized in fig. 3.7 as a function of the threshold parameter ϵ, both for the 2-mode and the 4-mode fiber that we are simulating, showing worst-case scenario, i.e. the maximum over the 5 nonlinear terms.

We can observe that for both fibers, using $\epsilon = 0.01$, only about 10 percent of the coefficients would not be discarded, resulting in a significant speedup of the simulations.

We ran a series of simulations to evaluate the effect of using this technique noting no significant changes, confirming the viability of the approach. For this reason, the results presented in this thesis are all obtained using a threshold parameter of $\epsilon = 0.01$.

3.4.3 *Simulation results*

With the theoretical and numerical models defined, the equations in eq. (3.39) can be integrated for the two different fibers that have

been previously introduced, aiming to understand how different levels of linear coupling can influence the statistics of the distributed amplifier's gain. For the specified levels of perturbation length L_κ and correlation length L_c, we solve eq. (3.39) for an ensemble of N_e different realizations of the process describing the orientation angle $\theta_k(z)$ of the perturbations acting on the fiber, obtaining the evolution of the vector of complex amplitudes $\mathbf{a}(z;\omega_s)$ for each realization.

Remembering the way in which the entries of $\mathbf{a}$ are organized, i.e. with alternating x- and y- polarizations of a mode, the total power in each spatial mode at the signal frequency ω_s is given by

$$P_\nu(z) = |a_{2\nu}(z)|^2 + |a_{2\nu+1}(z)|^2, \quad \nu \in [0, \dots, N_s - 1], \tag{3.100}$$

where N_s is the number of spatial modes, which is equal to 3 for the 2-mode fiber and to 6 for the 4-mode fiber. The amplifier gain for mode ν is then computed for the kth realization of the process $\theta_k(z)$ as

$$G_\nu(k) = \frac{P_\nu(L;\theta_k)}{P_\nu(0;\theta_k)}, \tag{3.101}$$

computing its ensemble mean as

$$\langle G_\nu \rangle = \frac{1}{N_e} \sum_{k=1}^{N_e} G_\nu(k). \tag{3.102}$$

From the same quantities, the normalized signal power variance at the end of the amplifier can also be evaluated as [111]

$$\sigma_\nu^2 = \frac{\langle P_\nu^2(L;\theta_k)) \rangle}{\langle P_\nu(L;\theta_k) \rangle^2} - 1, \tag{3.103}$$

where $\langle \cdot \rangle$ indicates the mean over the ensemble of perturbation realizations, as in eq. (3.102).

Regarding the simulation parameters, we set to 0 the coefficient σ related to Kerr effects, in order to focus only on Raman-related phenomena. To this end, we set the peak Raman gain coefficient to $g_R = 1 \times 10^{-13}\,\mathrm{m\,W^{-1}}$, and use eq. (3.59) to obtain the corresponding values of the response functions. We consider no MDL, and set the fiber loss coefficient to $\alpha_s = \alpha_p = 0.2\,\mathrm{dB\,km^{-1}}$. Recalling the considerations made in section 3.4.2, a threshold parameter of $\epsilon = 10^{-2}$ was set to optimize the computation of the nonlinear part of the equations in the split-step algorithm, whose step size is set to $\delta z = L_c/10$. The information-bearing signals have an initial power of $-40\,\mathrm{dBm}$ per spatial mode, while a total of 1 W is launched on the pump wavelength, equally distributed among the supported modes.

The length of the simulated fibers is set to 50 km, which has been verified to be sufficient for the amplification dynamics to be fully exhausted for our simulation conditions.

The metrics in eq. (3.102) are evaluated for two different initial conditions on the polarization of both pump and signal modes. In the first case, the polarization of each spatial mode at the start of the fiber for the signal frequency is set to be linear and parallel to that of the modes of the pump, with the corresponding sub-vector of complex amplitudes at $z = 0$ reading

$$\mathbf{a}_\nu(0;\omega_s) = \sqrt{P_\nu(0;\omega_s)}\begin{pmatrix}1\\0\end{pmatrix} \quad \mathbf{a}_\nu(0;\omega_p) = \sqrt{P_\nu(0;\omega_p)}\begin{pmatrix}1\\0\end{pmatrix} \tag{3.104}$$

On the contrary, the second case considers orthogonal linear polarizations between signal and pump, defining the following sub-vectors at the start of the fiber

$$\mathbf{a}_\nu(0;\omega_s) = \sqrt{P_\nu(0;\omega_s)}\begin{pmatrix}1\\0\end{pmatrix} \quad \mathbf{a}_\nu(0;\omega_p) = \sqrt{P_\nu(0;\omega_p)}\begin{pmatrix}0\\1\end{pmatrix}. \tag{3.105}$$

3.4.3.1 *2-mode fiber*

Starting with the simulated 2-mode fiber, the ensemble average $\langle G_\nu \rangle$ for these two cases is represented as a function of the perturbation beat length L_κ in fig. 3.8, indicated by the symbols $\parallel$ and $\perp$, respectively. Correlation lengths of 10 m (left) and 100 m (right) are considered. The number of simulated fibers for each combination of parameters is set to $N_e = 5000$.

Focusing on the case of $L_c = 10\,\mathrm{m}$, we can clearly discern three regimes of propagation for the given fiber length, similarly to the case of randomly-birefringent SMFs [112]. For low degrees of birefringence, i.e. $L_\kappa \gg 1$, the mean gain is maximized on each mode for parallel input polarizations, reaching $\approx 41\,\mathrm{dB}$ for the LP_{11} modes, and $\approx 37\,\mathrm{dB}$ for the LP_{01} mode. The higher gain experienced by the LP_{11} with respect to the LP_{01} is explained by the fact that the higher order mode group is pumped with twice the power of the fundamental one due to its 4-fold degeneracy. The difference in pumping power, together with the efficient nonlinear interaction between the two degenerate spatial modes forming the LP_{11} group , counterbalances its larger effective area (lower nonlinear overalap integral). In this regime, with orthogonal input polarizations, the gain is instead minimized, almost totaling the link-losses of $0.2\,\mathrm{dB\,km^{-1}} \times 50\,\mathrm{km} = -10\,\mathrm{dB}$; it is not exactly equal since a small gain exists for orthogonally-polarized pumps, as seen in section 3.2. In this regime, the only interaction

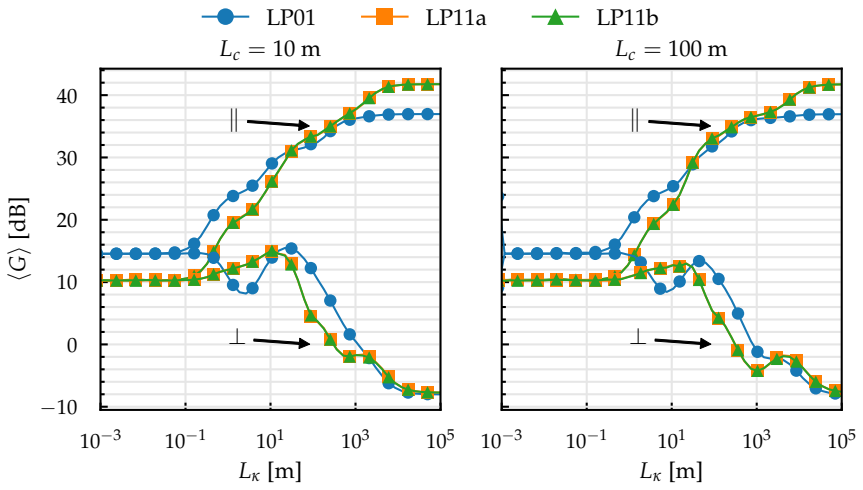


Figure 3.8: Mean gain of the spatial modes of the considered 2-mode fiber as a function of the coupling beat length, for two values of correlation length. Represented are the cases when signal and pump polarization are parallel (indicated with ∥), and when the two are orthogonal (indicated with ⊥), at the start of the fiber.

.

between the three modes is determined by the nonlinear coupling through the overlap integrals $\mathcal{Q}_{\nu\mu\rho\eta}$. The relative angle between the pump and signal polarization remains almost constant along the fiber due to the low values of coupling, explaining the observed behavior.

For intermediate birefringence values, when pump and signal polarizations are aligned at the input, the resulting gain quickly decreases, as the effect of linear coupling begins to scramble their relative orientation, reducing the Raman gain efficiency. It is interesting to note the presence of several inflection points that are present in this regime for all three modes, although they are more accentuated for the case of the LP_{11} modes. The opposite is true for the case of orthogonal polarizations, where the effect of linear coupling is beneficial to increase the amplifier gain. In this condition, similarly to the SMF case, the gain for all three modes first reaches a maximum value before decreasing again.

In the high-birefringence regime instead, for $L_\kappa \ll 1$, the alignment between pump and signal polarization is quickly lost due to the rapid exchange of power inside the group of degenerate modes, and the gain converges to the same average value both for the parallel and orthogonal input polarizations. Here we notice the first difference with respect to the case of SMFs, where the amplifier gain in high PMD conditions is approximately half (in dB) of that obtained with no PMD. In fact, the gain of both LP_{11a} and LP_{11b} modes is reduced to just 10 dB,

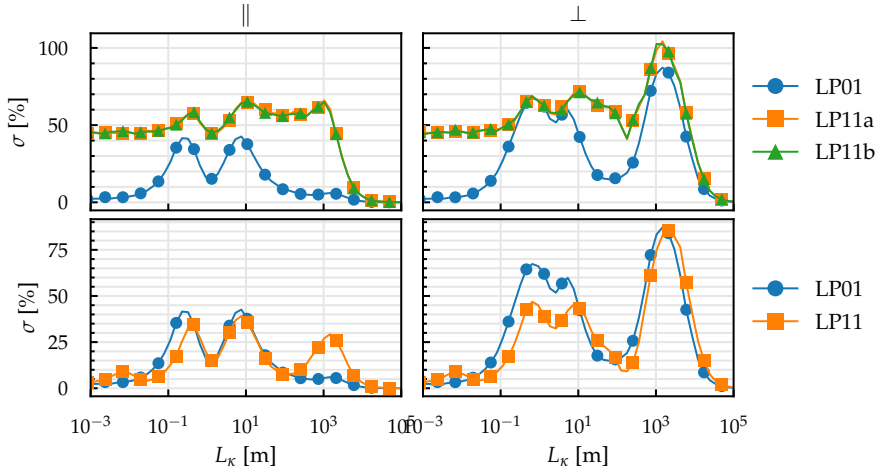


Figure 3.9: Signal power variation of the gain of the 2-mode amplifier as a function of the coupling beat length, for $L_c = 100\,\mathrm{m}$, for parallel (left column, $\parallel$) and orthogonal (right column, $\perp$) relative polarization orientation between pump and signal. Top row: power variation of the individual modes. Bottom row: power variation on the mode groups.

resulting in the reduction of the equivalent gain coefficient of about 4, while the LP_{01} mode sees its average gain reduced to approximately 14 dB. Increasing the correlation length to 100 m, the dynamics remain unaltered, but shifted toward higher beat length values.

At first sight it might seems that the MDG on modes of the LP_{11} group is zero for any value of birefringence, essentially confirming the results of [57] stating that the Raman gain can be considered equalized among the quasi-degenerate modes of each group.

Further insights can be gained by analyzing the signal power fluctuations through the evaluation of its standard deviation, as in eq. (3.103). In fig. 3.9, its value is represented as a percentage, and depicted for each mode as a function of the coupling beat length for the parallel (top left) and orthogonal (top right) polarization case, and for a correlation length of $L_c = 100\,\mathrm{m}$.

We can observe two different behaviors, depending on the modes we consider. Starting from the LP_{01} mode, for parallel polarizations, the three birefringence regimes highlighted previously for the amplifier gain are clearly recognizable. For extreme values of beat lengths, the signal variation goes to 0, resulting in stable output power and a clearly defined gain. On the other hand, for intermediate values of L_κ, two clearly defined peaks reaching a value of approximately 45 % appear. This is consistent with what has been demonstrated in the case of SMFs, albeit in the counterpumping configuration.

Regarding the modes belonging to the LP_{11} group, a similar behavior is found for long beat lengths, where the signal fluctuations are negligible. Differently, for intermediate values of birefringence, the standard variation is fairly higher than the LP_{01} mode, reaching 70%, and showing a local maximum for much longer beat lengths. Additionally, in the high birefringence regime, the fluctuations do not decrease to 0 as for the fundamental mode, but instead remain constant at approximately 50%. This behavior can be intuitively explained by the fact that in the LP_{11} group is 4-fold degenerate, meaning that the power exchange occurs both between polarizations and spatial modes, as seen in the analysis of its linear coupling matrices. This means that the LP_{11a} and LP_{11b} modes act as the individual polarizations of e.g. the fundamental mode of a SMF system, in which PMD makes the Stokes vector of the amplified signal to randomly explore the Poincaré sphere, varying the power in each polarization depending on the specific realization of the birefringence process. In order to verify this, we compute eq. (3.103) on the total power contained in the LP_{11} group, and illustrate the results for parallel input polarizations in the bottom left graph of fig. 3.9. A similar behavior to that of the fundamental mode is now obtained, with the group power variation going to 0 for short beat lengths. This plot also emphasizes the presence of two secondary peaks at approximately $L_\kappa = 10^{-2.5}$ m and $L_\kappa = 10^3$ m, with the latter being significantly more prominent with a peak value of 30%. Interestingly, for the values of L_κ near this peak, a small increase in the signal variance is observed for the LP_{01} mode, which, in absence of inter-group coupling, can be explained by the Raman interaction between the two groups through the overlap integrals Q. The location of these peaks correspond to that of the inflection points observed in the behavior of the average gain, in fig. 3.8. The same considerations can be made for the case of orthogonal input polarizations from the top right and bottom right graphs of fig. 3.9, showing the same dynamics, although much more accentuated.

These results show that the gain equalization inside each mode group is only obtained in the average sense, since random coupling caused by residual stress birefringence and core ellipticity of the fiber can make the power on each spatial mode fluctuate significantly for realistic values of L_κ. The equalization of the LP_{11} mode reached for $L_\kappa \gg 1$ is also, evidently, not resulting from coupling, but can be instead explained the symmetry of the nonlinear coefficients Q.

Moreover, the total power of the entire group only shows negligible fluctuations for really low values of beat length, observed typically in polarization-maintaining FMFs [133].

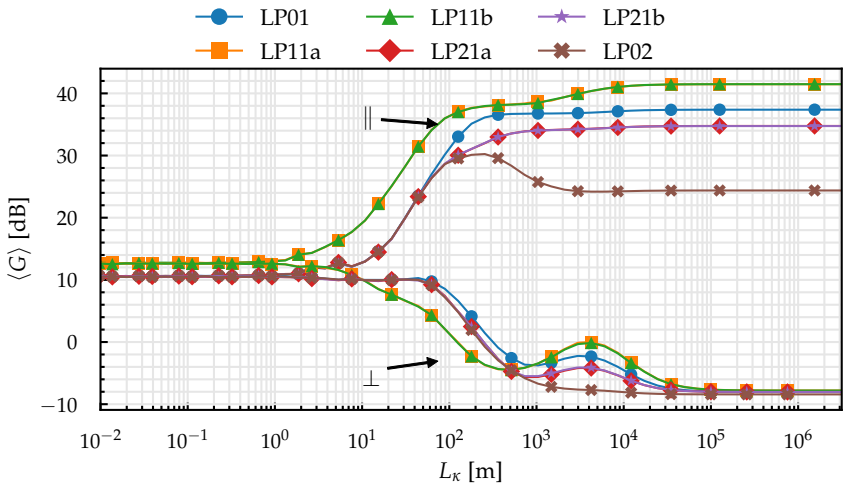


Figure 3.10: Mean gain of the spatial modes of the considered 4-mode fiber as a function of the coupling beat length, for $L_c = 100\,\mathrm{m}$. Represented are the cases when signal and pump polarization are parallel (indicated with $\parallel$), and when the two are orthogonal (indicated with $\perp$).

3.4.3.2 *4-mode fiber*

Using the models for birefringence and core-ellipticity detailed in section 3.4.1, and from the results of the previous sections, we saw that no linear inter-group effects are present for a 2-mode fiber. When considering a fiber that supports 4 groups of LP modes instead, inter-group coupling is expected between the LP_{01}, LP_{02}, and LP_{21} groups, as a result of the core-ellipticity perturbation tensor. Modes of the LP_{11} group instead only experiences intra-group coupling, meaning that similar behaviors to the 2-mode fiber are expected for this group. Employing the same simulation conditions detailed in the previous section, the average gain and signal power fluctuations are obtained using an ensemble of $N_e = 5000$ realizations of the perturbation process. Here, the analysis is limited to the case in which the correlation length is $L_c = 100\,\mathrm{m}$ since by changing its value equivalent results are obtained, provided that the perturbation beat length is properly scaled, as seen from the results for the 2-mode fiber.

The mean gain $\langle G_\nu \rangle$ is visualized in fig. 3.10 as a function of the beat length L_κ, considering parallel (indicated with $\parallel$) and orthogonal (indicated with $\perp$) input polarizations between pump and signals.

Similar results to the 2-mode fiber are found in the strong birefringence regime: for parallel polarizations, the gain is maximized, with LP_{11} modes experiencing the highest amplification of approximately 42 dB and the LP_{02} mode showing instead the smallest gain, at 24 dB.

LP_{01} and LP_{21} groups instead exhibit a maximum gain of 35 and 37 dB, respectively. For orthogonal input polarizations the gain is minimized, and equal to approximately -8 dB for each mode, showing a 2 dB reduction of the total link losses due to the minimal contribution of the cross-polarized Raman response on the amplification process.

Decreasing the beat length, the effect of inter-group coupling can be readily observed for parallel polarizations, and is most apparent on the behavior of the average gain of the LP_{02} mode, which quickly merges with the curves for the LP_{21} modes for beat length values in the order of 10^2 m. Soon after, the LP_{01} gain joins the aforementioned curves, becoming fully coupled for $L_\kappa \approx 10^{1.5}$ m and forming a "supergroup". This is in accordance with the computed coupling matrix for core-ellipticity in fig. 3.5 (b), where we can observe the slightly higher coupling coefficients between the LP_{02}-LP_{21} group pair with respect to the LP_{01}-LP_{21} combination. For typical values of birefringence $\delta_n \approx 10^{-7}$ for SMFs, corresponding to a beat length $L_\kappa \approx 10\,\text{m}$ following fig. 3.6, the effect of linear inter-group coupling has to be accounted for even when modest span lengths of 50 km are considered. When extrinsic perturbations such as twisting and bending, deriving mostly from cabling and installation, are considered, the effect of inter-group coupling on the Raman gain is expected to be enhanced, and valid for shorter spans under the same birefringence and core ellipticity parameters. Regarding the orthogonal polarizations case, the effect of inter-group coupling is similar, although the convergence between the LP_{21} and LP_{02} group pairs is reached for slightly longer beat lengths. The general behavior of reaching a local maximum gain observed in the 2-mode case is also present for all the modes except the LP_{02}, for which the increase is monotonic.

Finally, for high birefringence values corresponding to beat lengths $L_\kappa < 1\,\text{m}$, the curves for parallel and orthogonal polarizations converge to a common average gain of 10 dB and 13 dB for the LP_{01}-LP_{21}-LP_{02} supergroup and the LP_{11} group, respectively.

As a general remark, modes belonging to groups with 4-fold degeneracy, i.e. LP_{11} and LP_{21}, maintain the same average gain for the entire interval of considered beat lengths, regardless of the relative orientation of pump and signal polarizations, similarly to the case of the 2-mode fiber.

Next, the signal power fluctuation is evaluated for each mode with eq. (3.103) and illustrated in fig. 3.11 as a function of the beat length L_κ. The figure is organized similarly to the case of the 2-mode fiber, with the left and right columns corresponding to the parallel and orthogonal polarizations cases, indicated with $\parallel$ and $\perp$, respectively; the top row corresponds to the power fluctuation evaluated on each individual spatial mode, while the same metric computed on the total

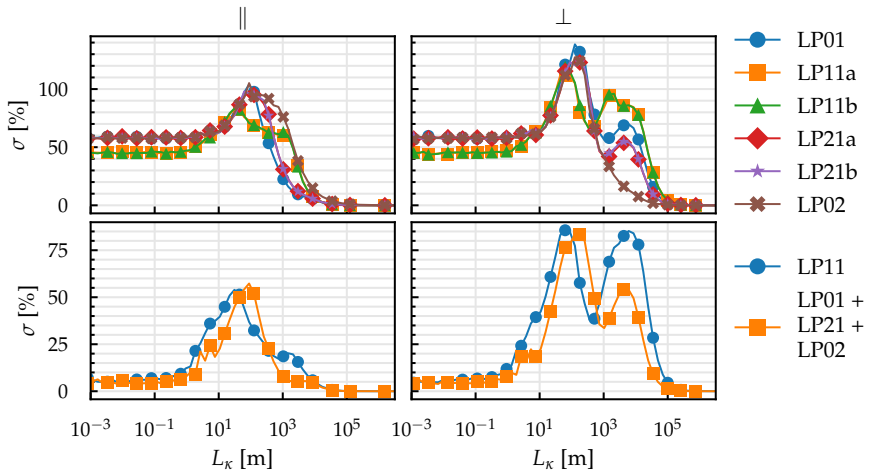


Figure 3.11: Signal power variation of the 4-mode amplifier as a function of the coupling beat length, for $L_c = 100$ m, for parallel (left column, ∥) and orthogonal (right column, ⊥) polarizations between pump and signal. Top row: power variation of the individual modes. Bottom row: power variation on the "supergroups".

.

power of the LP_{11} and the LP_{01}-LP_{21}-LP_{02} supergroup is depicted in the bottom row.

Starting from the parallel polarizations case, we can see from the top-left graph of fig. 3.11 that the signal power variation on each mode behaves similarly, presenting negligible fluctuations when the birefringence is unrealistically low, for $L_\kappa > 10^4$ m. Decreasing the beat length, the variance quickly increases up to a maximum of 100% when $L_\kappa \approx 10^{1.5}$ m, i.e. when the LP_{01}-LP_{21}-LP_{02} supergroup is formed. Interestingly, even though no inter-group coupling affects the LP_{11} group, its variance is increased with respect to the 2-mode fiber case, which can be explained by the nonlinear coupling occurring with each mode forming the supergroup. For low birefringence, the power fluctuation stabilizes to about 50% for the LP_{11} mode, and to 55% for the modes that constitute the supergroup. In this case, the high power fluctuation that we have so far encountered just for the 4-fold degenerate modes is also present for the $LP_{o,p}$ groups consisting of just the two polarization degeneracies, as a consequence of core ellipticity causing inter-group coupling between those modes, as seen in section 3.4.1.

If we consider the variance of the total power of the LP_{11} groups and LP_{01}-LP_{21}-LP_{02} supergroup, the high-fluctuations that are present in the high-birefringence regime quickly decay for beat lengths approaching $L_\kappa = 1$ m, as can be observed in the bottom-left graph of fig. 3.11. The highest signal variation is still obtained at $L_\kappa \approx 10^{1.5}$ m for both

supergroups, each reaching fluctuations of about 55%. Comparing these results with those obtained for the 2-mode fibers, we can notice the absence of multiple prominent peaks; instead, a weak increase can be noticed at $L_\kappa \approx 10^3$ m for the LP_{11} group, where an inflection point is present in the corresponding gain curve in fig. 3.10.

These effects become enhanced when considering orthogonal input polarizations, as seen in top- and bottom- right graphs of fig. 3.11, where the peak we observed at $L_\kappa \approx 10^3$ m for parallel polarizations is dramatically more evident for modes belonging to both groups. Considering the variation on the individual spatial modes, the fluctuation increases with the average gain of the respective mode, as seen in fig. 3.10.

3.5 CONCLUSIONS

In this chapter, a theoretical and numerical model that describes the propagation dynamics of the modes on two different wavelengths of a distributed Raman-amplified FMF links has been derived. Using analytical models present in the literature, the effect of intrinsic perturbations such as stress birefringence and core ellipticity, originating from the fiber manufacturing process, have been used to derive the linear coupling coefficients for two FMFs, supporting 2 and 4 groups of LP modes, respectively. The statistics on the gain of the amplifier have been gathered by simulating thousands of realizations of the stochastic linear coupling process described by the angle $\theta(z)$ along which the birefringence and core ellipticity perturbations are aligned.

For the 2-mode fiber, the considered perturbations induce coupling among the quasi-degenerate modes belonging to the same LP group. Interaction between the two groups only occur through the nonlinear coupling coefficients that determine the Raman amplification process. Three coupling regimes are identified, extending the results of SMF systems. When mode coupling is weak or negligible, the average gain on each mode is maximized (minimized) when signal and pump polarizations are parallel (orthogonal) at the fiber input, with the output signals showing no power fluctuations. As coupling increases, the average gain starts to decrease (increase), and the power variation on each mode group is maximized. Finally, for high values of birefringence, the average gain on each mode group for parallel and orthogonal input polarizations reach the same value. In this condition, the total power of each group of modes also shows reduced fluctuations.

When considering a FMF supporting the propagation of 4 groups of LP modes similar effects can be observed, with the addition of core ellipticity causing inter-group coupling between the LP_{01}, LP_{21}, and LP_{02} groups. For typical values of birefringence, the average gain

of each mode belonging to this "supergroup" converges to the same value for a 50 km-long fiber. At the same time, in this regime, high fluctuations of the signal power can be observed.

In general, in the strong-coupling regime characterized by high birefringence values, the fluctuations of the total power of the group of coupled modes tends to zero, but the power on each spatial mode keeps exhibiting high variance, behaving similarly to the individual polarization components in a SMF system, in which PMD causes the Stokes vector of the fundamental mode to rotate on the Poincaré sphere and causing a rapid exchange of power between each orthogonal polarization states. Furthermore, this regime is unlikely to be reached in a single span, with practical values of birefringence corresponding to beat lengths in the order of tens of meters, becoming relevant only when considering long-haul multi-span transmissions.

Part II

SUPERCONTINUUM GENERATION

4

PRINCIPLES OF SUPERCONTINUUM GENERATION

4.1 INTRODUCTION

As introduced in chapter 1, when a short, high-power optical pulse propagates through a medium, the interplay between its dispersive and nonlinear properties contributes to a massive broadening of the initial pulse spectrum, forming what is known as supercontinuum (SC). Spectral broadening and the generation of new frequency components is an inherent feature of nonlinear optics, and has been studied in bulk materials since the 1960s after the discovery of the laser. In fact, in just a few years, from 1961 to 1966, second harmonic generation, frequency mixing, parametric generation, third harmonic generation, four-wave mixing (FWM), self-phase modulation (SPM), self-focusing, stimulated Brillouin scattering (SBS), and stimulated Raman scattering (SRS) were all characterized in a variety of different materials [134].

However, it was not until the early 1970s, thanks to the seminal work by Robert R. Alfano and Stuart L. Shapiro [135], that mechanisms to generate an extended "white light" source where first demonstrated and later applied for time resolved spectroscopy using picosecond pulses in a borosilicate glass sample [136].

During these years of early research in the field, this phenomenon was commonly referred to as frequency broadening, anomalous frequency broadening, or white light generation; the first use of the term "supercontinuum" is attributed to Gersten, Alfano, and Belic after describing the theory of the interaction between SPM and SRS [137]. Thanks to the development of low-loss optical fibers, the same effects characterized previously in bulk media were soon demonstrated in silica fibers at much lower power levels. Optical fibers then became the preferred medium to study nonlinear optical effects, leading to the birth of the field of nonlinear fiber optics. From 1972 to 1978, most of the effects that play a major role in supercontinuum generation in optical fibers were demonstrated, including SRS, SBS, the optical Kerr effect, FWM, SPM [134].

Optical soliton generation, a key process in supercontinuum generation in fibers, arising from the balance between SPM and anomalous dispersion, was first theorized in optical fibers in 1973 [138] and experimentally demonstrated only in 1980 [139] due to technological challenges involved in developing sources capable of delivering picosecond pulses in the anomalous dispersion regime of silica-based

fibers, near 1.3 µm. When the fiber pumped in this region, these unperturbed sub-picosecond pulses forming in the fiber propagate with high peak power, barely affected by pulse broadening due to dispersion, and interact with SRS and FWM to produce a continuum extending to the long-wavelength region of the spectrum [140].

Even though theoretically solitons form under strict conditions on the input pump pulse, it was demonstrated that any reasonable pulse shape will evolve into a soliton [141]. The energy that is not required to form solitons, in such cases, is "shed" into what are known as "dispersive waves", which are mainly responsible for short-wavelength generation [142].

For a given pump wavelength the generated supercontinuum is critically dependent on the nonlinearity and the dispersion. With the advent of photonic-crystal fibers (PCFs), however, it was shown that by adjusting the size of the air holes forming the photonic crystal and their relative distance, control of the dispersion profile was possible [134]. Fibers could then be manufactured with their zero dispersion wavelength (ZDW) tuned to the available high power sources, in order to take full advantage of the soliton dynamics for the generation of SC.

Furthermore, PCFs supporting the propagation of a single mode for any wavelength were demonstrated, improving the output beam characteristics compared to conventional fibers due to precise control of the dispersion profile and avoiding inter-modal effects [143]. In addition, much higher nonlinear coefficients derived from a reduced effective mode field diameter could be achieved compared to conventional fibers [144].

Combining the dispersion tailoring capabilities of PCFs with advances in fiber manufacturing process using materials with different transmission windows than silica, such as tellurite, fluoride, or chalcogenide, SC multiple-octave SC sources could be developed from the visible [145] to the mid-infrared portion of the spectrum [146].

This flexibility, coupled with their extremely high brightness and broad bandwidths, made SC sources a candidate for several applications, such as optical coherence tomography [147] and optical frequency metrology [148], where a high degree of temporal coherence is needed.

However, as detailed in [149], the soliton-driven dynamics that generate SC by pumping in the anomalous dispersion region of the fiber cause the temporal breakup of the injected pulse and a high sensitivity to noise, which translates to high fluctuations of the output spectrum and therefore loss of coherence.

A solution was proposed in [150], in the form of all-normal dispersion (ANDi) PCFs, which are designed to have a flattened convex

profile of normal dispersion. By pumping the fiber near its minimum dispersion wavelength (MDW) the pulse broadening is minimized, and noise-sensitive soliton dynamics are avoided, with spectral broadening being mainly associated to highly coherent SPM and FWM processes. As a consequence, low noise-sensitivity and preservation of the input pulse can be achieved, with uniform spectral and temporal profiles [31].

In the following sections, the different linear and nonlinear phenomena that are at the base of SC generation in optical fibers are reviewed. A comparison between the conventional anomalous dispersion-based SC and the ANDi SC regime is also given, highlighting the dynamics at play.

The content of this chapter serves as a review of the physical processes driving the phenomenon of SC, and is based on Refs. [26, 31, 134].

4.2 FUNDAMENTAL PROCESSES FOR SC GENERATION

In the simplest case in which a single optical pulse, linearly polarized at the carrier frequency ω_0, is launched in the fiber as to excite only a single mode, it can be shown [151] that its slowly varying complex envelope $A(z,t)$ satisfies the following propagation equation

$$\frac{\partial A}{\partial z} + \frac{\alpha}{2}A = i\sum_{m=1}^{\infty}\frac{i^m\beta_m}{m!}\frac{\partial^m A}{\partial t^m} + i\gamma\left(1+\frac{i}{\omega_0}\frac{\partial}{\partial t}\right) \tag{4.1}$$

$$\times\left[\int R(t')\left|A(z,t-t')\right|^2 dt'\right]A(z,t) \tag{4.2}$$

where α is the usual attenuation coefficient, γ is the nonlinear parameter defined as

$$\gamma = \frac{\omega_0 n_2(\omega_0)}{c_0 A_{\text{eff}}} \quad A_{\text{eff}} = \frac{\left[\iint F(x,y,\omega_0)^2 dxdy\right]^2}{\iint |F(x,y,\omega_0)|^4 dxdy}, \tag{4.3}$$

where n_2 is the nonlinear refractive index of the considered fiber, A_{eff} is its fundamental mode's effective area, c_0 is the speed of light in vacuum, and the β_m terms are the Taylor expansion coefficients of the frequency-dependent propagation constant around the central frequency ω_0

$$\beta(\omega) = \beta_0 + (\omega-\omega_0)\beta_1 + \frac{1}{2}(\omega-\omega_0)^2\beta_2 + \ldots, \quad \beta_m = \left(\frac{d^m\beta}{d\omega^m}\right)_{\omega=\omega_0}. \tag{4.4}$$

The integral in eq. (4.1) contains the nonlinear electronic and nuclear response of the material, contributing, respectively, to the Kerr and Raman effects, as detailed in chapter 3. When considering pulse propagation dynamics, it is usually decomposed in the following way

$$R(t) = (1 - f_R)\delta(t) + f_R h_R(t), \tag{4.5}$$

where f_R is the fractional Raman contribution to the overall third order nonlinearity.

As seen in chapters 2 and 3, the form of the response function can either be obtained experimentally by measuring the Raman gain of the material and using the Kramers-Kronig relation [152], or analytically approximated as in [59, 153].

Although eq. (4.1) must be solved when considering the propagation of ultrashort pulses, it can be simplified considerably pulses large than a few picoseconds are considered. In fact, the Raman response function $h_R(t)$ has an appreciable magnitude only for $t < 1$ ps, and can be approximated with $\delta(t)$ for pulses wider than 3 to 4 ps. The derivative term containing ω_0, and dispersive terms higher than the third order can be ignored, leading to the following equation:

$$\frac{\partial A}{\partial z} + \frac{\alpha}{2}A + \beta_1\frac{\partial A}{\partial z} + \frac{i\beta_2}{2}\frac{\partial^2 A}{\partial t^2} - \frac{\beta_3}{6}\frac{\partial^3 A}{\partial t^3} = i\gamma|A|^2 A. \tag{4.6}$$

When losses and the third-order dispersion are neglected ($\alpha = 0$, $\beta_3 = 0$), the equation is reduced to the familiar Nonlinear Schrödinger Equation (NLSE) [151].

It is useful to introduced some normalized variables. Any input pulse injected in a fiber has its amplitude described by $A(0,t) = \sqrt{P_0}S(t/T_0)$, where P_0 is the peak pulse power, $S(t)$ describes the pulse shape, and T_0 is a measure of the pulse width.

For a fiber of length L, the following normalized variables are introduced

$$Z = z/L, \qquad \tau = (t - \beta_1 z)/T_0, \qquad A = \sqrt{P_0}, \exp(\alpha z/2)U. \tag{4.7}$$

defining a reference time frame moving with the pulse at the group velocity, and including the exponential decay of the pulse amplitude in the definition of the normalized amplitude U. Equation (4.6) then assumes the form

$$\frac{\partial U}{\partial Z} + \frac{isL}{2L_D}\frac{\partial^2 U}{\partial \tau^2} - \frac{\beta_3}{6T_0^3}\frac{\partial^3 U}{\partial \tau^3} == \frac{iL}{L_{NL}}e^{-\alpha z}|U|^2 U,, \tag{4.8}$$

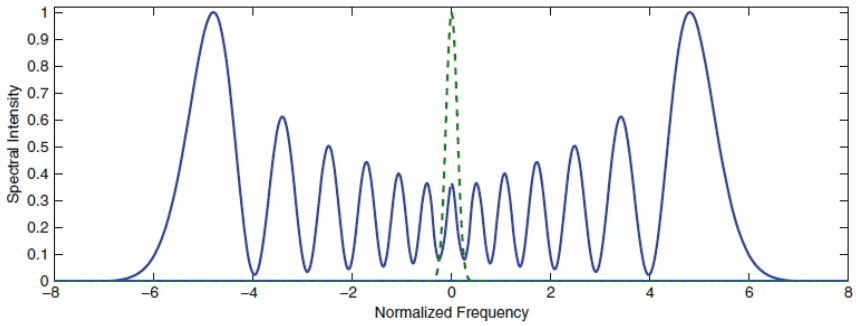


Figure 4.1: Effect of SPM on the spectrum of a Gaussian pulse for $\phi_{max} = 40$ [26].

where $s = \text{sign}(\beta_2)$, and the dispersion and nonlinear lengths, defined as

$$L_D = \frac{T_0^2}{|\beta_2|}, \qquad L_{NL} = \frac{1}{\gamma P_0}. \tag{4.9}$$

provide the length scales over which dispersive and nonlinear effects become important.

The ratio between the two quantities can be used to determine if either dispersion or nonlinear effects dominate each other. Defining

$$N^2 = \frac{L_D}{L_{NL}} = \frac{\gamma P_0 T_0^2}{|\beta_2|}, \tag{4.10}$$

nonlinear effects become negligible when the fiber and the pulse parameters are such that $N^2 \ll 1$, and only the well-known dispersive pulse-distortion dynamics are considered.

4.2.1 *Self-phase modulation*

When the fiber length L is such that $L \ll L_{LD}$ but $L \gg L_{NL}$, the dispersion terms become negligible with respect to the nonlinear term. In this case, pulse propagation is dominated by SPM, which, when acting alone, produces changes to the pulse spectrum but leaves its temporal profile intact.

When $N \gg 1$, eq. (4.8) has a closed form solution which reads

$$U(L,\tau) = U(0,\tau)\exp\left(i\phi_{NL}(L,\tau)\right), \tag{4.11}$$

where $\phi_{NL}(L,\tau) = |U(0,\tau)|^2(L_{eff}/L_{NL})$, and $L_{eff} = [1-\exp(-\alpha L)]/L$ is the fiber effective length.

The time dependent phase ϕ_{NL} induces a chirp on the pulse, creating new frequency components as the pulse propagates in the fiber,

broadening its spectrum. The maximum phase shift ϕ_{max} occurs at the pulse center, and is given by

$$\phi_{max} = \frac{L_{eff}}{L_{NL}} = \gamma P_0 L_{eff}. \tag{4.12}$$

An example is given in fig. 4.1, where the original (dashed green) Gaussian input pulse spectrum and the effect of SPM (solid blue), are illustrated using $\gamma = 2\,\mathrm{W\,km^{-1}}$, $L_{eff} = 20\,\mathrm{km}$, and $\phi_{max} = 40$ at $P_0 = 1\,\mathrm{W}$.

4.2.2 *Soliton generation and fission*

When the dispersion and nonlinear lengths are shorter or comparable to the fiber length, dispersion and nonlinearity act together. When third-order dispersion is negligible, $\beta_3 = 0$, and the distance is normalized as $\xi = z/L_D$, eq. (4.8) becomes

$$\frac{\partial U}{\partial \xi} + \frac{is}{2}\frac{\partial^2 U}{\partial \tau^2} = N^2 e^{-\alpha z}|U|^2|U|. \tag{4.13}$$

In the anomalous dispersion case, it has been shown by Shabat and Zakharov in their seminal paper that any system described by the NLSE admits exact solutions in the form of solitons [154].

Under appropriate conditions, the balancing between dispersion and SPM enables the spatial, temporal, and spectral preservation of the propagating pulse. This happens when the initial pulse has shape $U(0,\tau) = \mathrm{sech}(\tau)$, with peak power P_0 and width T_0 such that $N = 1$. If the loss coefficient of the fiber α is sufficiently small, the soliton propagates unperturbed over long distances. The value of the parameter N determines the order of the soliton; the soliton that is created for $N = 1$ is called "fundamental".

For integer values of N, higher order solitons are formed in the fiber; instead of maintaining their temporal and spectral shape during propagation, these solitons evolve in a periodical fashion. High order solitons are, in fact, a nonlinear superposition of N fundamental solitons, and their periodic interference can be efficiently used to generate ultrashort pulses of only a few optical cycles, and is at the basis of the soliton laser [26].

This behavior is caused by the fact that the fundamental solitons forming the higher order soliton are propagating at the same speed inside the fiber. When third-order dispersion is non negligible, their propagation speed becomes slightly different and the high-order soliton breaks up into N individual solitons with their own peak power and temporal width in a phenomenon known as "soliton fission". Ad-

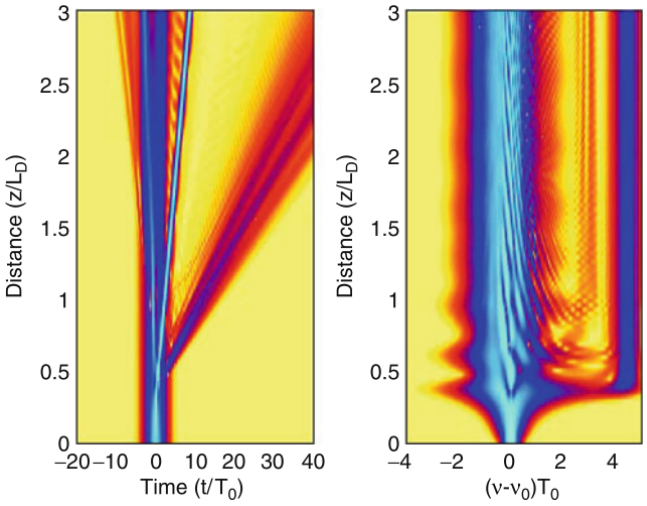


Figure 4.2: Temporal (left) and (spectral) evolution of the fission process of a high-order soliton($N = 3$) caused by third-order dispersion. [26].

ditionally, during the fission process each individual soliton sheds some of its energy in the form of dispersive waves. Due to phase-matching considerations, these waves are blue-shifted with respect to the original soliton, and are equivalent to the blue light emitted from a Cherenkov radiation process [149]. This effect is clearly visible in fig. 4.2, where the temporal (left) and spectral (right) evolution of the fission process of a third-order soliton is illustrated when setting $\delta_3 = \beta_3 L_D / 6T_0^3 = 0.02$, $N = 3$, and using a sech input pulse.

After an initial phase in which the injected pulse compresses due to the interference process between the three fundamental solitons, the pulse breaks up and the individual soliton propagate with a slightly different propagation velocity. At the same time, the high frequency dispersive waves, identified on the right image as the vertical blue line, rapidly lag behind the solitons.

Although from theory only hyperbolic secant pulses with specific durations and peak powers can turn into soliton, it has however been known since the early years of soliton physics that any reasonably shaped pulse can evolve into a soliton, with their excess energy converted in a blue-shifted dispersive wave [141].

4.2.3 *Influence of Raman scattering*

Another effect that disturbs the propagation of higher order solitons and contributes to the soliton fission phenomenon is the presence of the Raman term in eq. (4.1). The Raman gain amplifies the low-frequency components of the pulse at the expense of the high-frequency content of the same pulse that overlap in time. After fission,

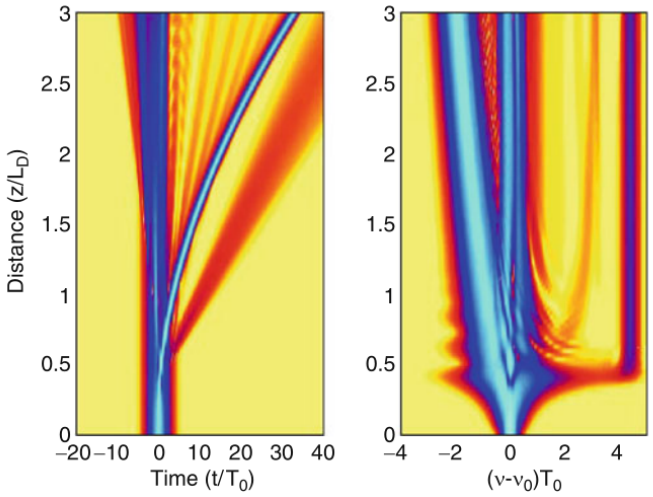


Figure 4.3: Fission of a third-order soliton and corresponding red-shift due to intrapulse Raman scattering. [26].

the spectrum of each fundamental soliton is shifted towards the longer wavelengths due to SRS, inducing what is known as Raman-induced frequency shift (RIFS). Its effect on the soliton fission process is shown in fig. 4.3 under the same conditions as fig. 4.2.

The higher peak power soliton, corresponding to a shorter pulse width, suffers a considerable shift towards the longer wavelengths due to SRS, and its temporal characteristics highlight a rapid slowdown caused by the change in group velocity during its continuous red-shift.

4.2.4 *Cross-phase modulation, four-wave mixing, and trapping of dispersive waves*

Although the combination of SPM, anomalous dispersion, soliton dynamics, and SRS causes a significant broadening of the injected pulse's spectrum, the example shown in fig. 4.3 is still not a continuum.

When increasing the soliton order by injecting higher power or shorter pulses, considering the derivative term of the nonlinear response of eq. (4.1), responsible of self-steepening, and including higher-order dispersion terms, new dynamics come into play.

After soliton fission occurs, the frequency-shifting solitons and the higher-frequency waves propagating in the normal dispersion region of the spectrum can temporally overlap and phase-match through cross-phase modulation (XPM). The interaction between short- and long- wavelength components of the spectrum is a consequence of the "U" shape of the group-velocity (or group-index) profile stemming from the typical dispersion curves of the fibers used for SC generation.

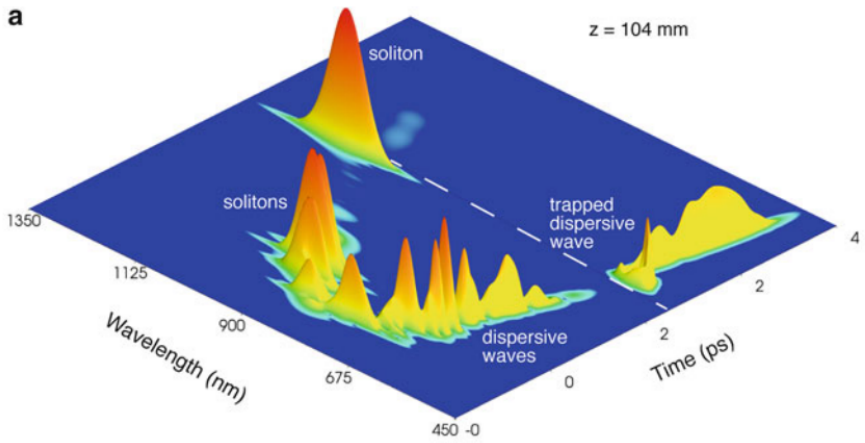


Figure 4.4: Spectrogram of the output supercontinuum generated by a 50 fs, 10 kW peak power pulse at 835 nm launched in a fiber with a ZDW of 780 nm. Adapted from [26].

The soliton RIFS consequently causes the blue-shift of the dispersive wave. This effect can be exploited by carefully designing the fiber dispersion and extend the generated SC on the short-wavelength side of the spectrum [155].

In fact, dispersive waves and solitons become "trapped" to each other, and propagate together in the fiber. This mechanism has been thoroughly studied, and can be explained by considering that initially the dispersive wave has a slower group velocity than the soliton, but the latter is continuously slowed down by the RIFS and will eventually collide. In the collision, the dispersive wave is reflected backward, i.e. with a slower group velocity and corresponding blue-shift [134]. This process then repeats as a consequence of the continuous RIFS of the soliton until the phase-matching cannot occur, for example when a second ZDW is encountered by the soliton, or due to losses.

Moreover, increasing the soliton order N enriches the generated SC, and new spectral components can arise due to XPM and FWM [151].

An example is shown in fig. 4.4, depicting the spectrogram of the generated SC when $N \approx 8.6$, obtained with 50 fs, 10 kW peak power pulses centered at 835 nm are injected in a PCF with ZDW of 780 nm. The U-shape of the group-velocity profile is clearly shown, with the central SPM-generated components arriving first. The red-shifting solitons and corresponding blue-shifting trapped dispersive waves are also indicated. The complex spectral and temporal characteristics of the output light are self evident.

4.3 ALL-NORMAL DISPERSION FIBERS

Although SC sources have revolutionized the world of nonlinear fiber optics and found widespread application in different scientific areas, their use in noise-sensitive or ultrafast photonics applications such as time-resolved spectroscopy, in which broad spectral bandwidths and coherent ultrashort pulses are required, is still not as common [31].

In fact, as seen in the previous section, the complex soliton dynamics that arise when a PCF is pumped in the anomalous dispersion region contributes to the temporal breakup of the output pulse, which is often comprised of several bursts of individual solitons and lower-level pedestals caused by dispersive waves. While it theoretically possible to compress the complex temporal structure of conventional SC pulses in a single ultra-short pulse, it could not be practically demonstrated due to the fine structure of its spectrum and the group delay shape [31].

Additionally, the soliton fission process that is responsible for the SC generation has been shown to be highly sensitive to quantum noise and to technical noise of the pump, such as relative intensity noise (RIN) or time jittering [149].

These phenomena, especially when using picosecond pulses, contribute to incoherent FWM and modulation instability (MI) processes seeded by noise fluctuations, which result in high variability of the generated spectrum and the respective temporal characteristics between subsequent output pulses, thereby deteriorating the coherence of the SC [149].

Although is has been known since the early days of SC generation that the temporal breakup of the pulse could be avoided when coherent SPM and FWM processes are stimulated by pumping the fibers in the normal dispersion region, the unavailability of short pulse sources resulted in incoherent SC caused by the Raman amplification of noise. When sub-picosecond sources became available, the interest in pumping fibers in the normal dispersion region faded due to the reduced spectral width caused by fast temporal broadening of the input pulse dictated by the steep dispersion profile of the available fibers. With the unprecedented flexibility that PCFs introduced in engineering the dispersion curve, all-normal dispersion (ANDi) fibers could be manufactured [150].

As their name suggests, these fibers are designed to have a dispersion curve entirely in the normal region, but close to zero and mostly flat near its minimum. An example is given in fig. 4.5, where the dispersion of a conventional commercially-available PCF with a ZDW at 780 nm is compared to that of an ANDi PCF, which is characterized by its minimum dispersion wavelength (MDW) at 1050 nm. By pumping

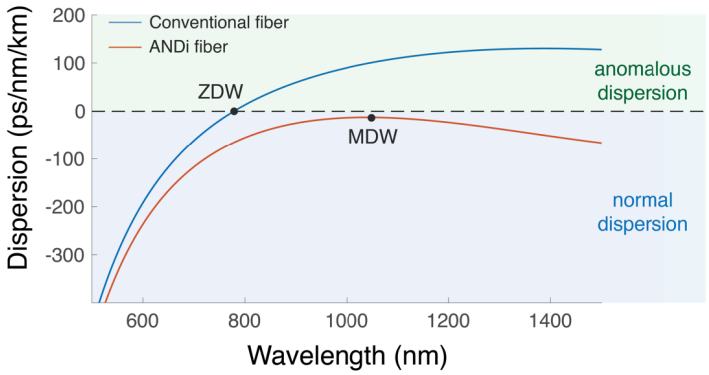


Figure 4.5: Typical PCF dispersion profiles for the conventional and ANDi SC generation regimes. The PCFs are preferably pumped on the anomalous dispersion region near their ZDW in conventional SC, or in the vicinity of the MDW for ANDi SC [31].

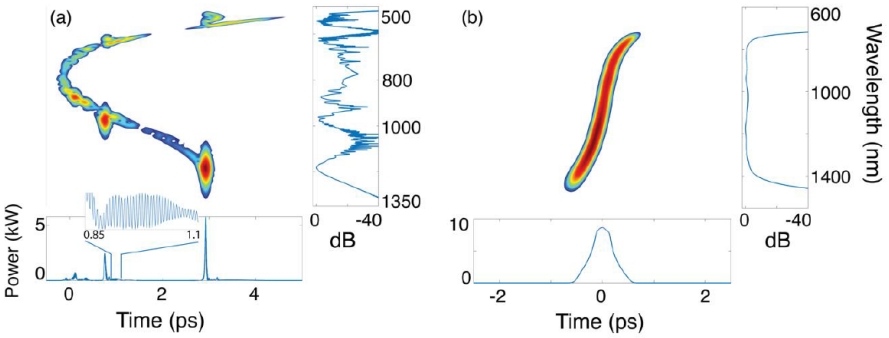


Figure 4.6: Projected axes spectrogram of (a) conventional SC and (b) ANDi SC after the broadening dynamics are concluded [31].

the ANDi fiber near its MDW, the temporal broadening of the input pulse is minimized and slowed down thanks to the flat dispersion profile in its vicinity.

The major differences in the obtainable SC between conventional and ANDi fibers can be understood by comparing their generated spectra and corresponding temporal characteristics. In fig. 4.6, the SC obtained by pumping the fibers in fig. 4.5 near their ZDW and MDW with 50 fs pulses of 10 kW and 90 kW peak power, respectively, are visualized in [31].

From the spectrogram of the conventional SC we can clearly distinguish the U-shape profile deriving from the group-velocity curve and the main features that characterize the soliton fission process. The corresponding spectrum shows the complex fine structures that are typical of conventional SC, while the pulse breakup is visible in the plot of the output pulse. On the other hand, the pulse shape is preserved in the case of the ANDi fiber, albeit with a much larger

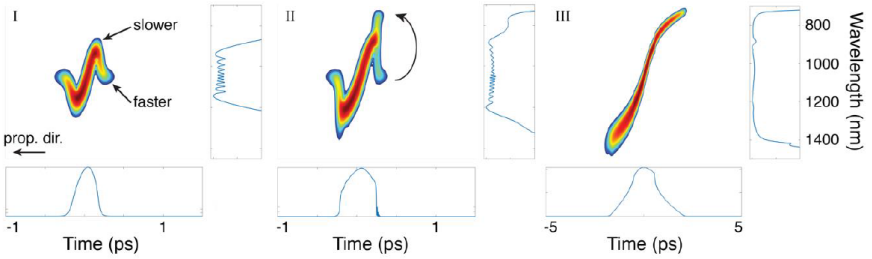


Figure 4.7: Dynamics of SC generation in ANDi fibers. Adapted from [31].

width compared to the individual solitons formed in the conventional SC. Correspondingly, the output spectrum shows similar width but a much smoother and flatter shape, with no evidence of distinct spectral features.

From its spectrogram we can observe a clearly defined distribution, partly due to group velocity strictly increasing with the wavelength. This almost-linear chirp greatly facilitates the compression of the pulse using spatial light modulators, enabling the generation of near transform-limited pulses narrower than 2 optical cycles and with higher peak power than the initial pump pulse [156]. From here it is intuitive to understand that having each wavelength located at a unique temporal position in the pulse is important to avoid the spectral or temporal interference that may cause irregularities either on the pulse shape or in its spectrum.

4.3.1 *Spectral broadening dynamics in ANDi fibers*

Contrary to conventional SC generation where the spectral broadening is mostly due to soliton fission and SRS, ANDi SC is mainly the result of SPM and FWM.

In fig. 4.7 the three main phases of the process of spectral broadening are highlighted.

In the initial phase (left), SPM is the dominating effect. Since the power is higher in the central part of the pulse, SPM broadening is more pronounced on the corresponding time interval; due to the normal dispersion, the longer generated wavelengths gain velocity while the shorter wavelengths slow down, giving the characteristic "S" shape to the pulse spectrogram. The presence of the same spectral components in different time instants generate the typical oscillations of SPM-broadened pulses.

While propagating, the group velocity dispersion accumulates and the trailing edge of the pulse, which has not been effected much by SPM due to the lower peak power, eventually catches up to the slower newly generated short wavelengths (middle). This has two main

consequence: first, the trailing pulse edge gets steeper, and second, the temporal overlap of two separate instantaneous frequencies leads two the generation of new frequency components via a degenerate FWM process called optical wave breaking (OWB). OWB then continuously transfer energy from the central part of the spectrum to its edge, contributing to the broadening phenomenon. The same process occurs also between the longer SPM-generated components and the leading pulse edge, generating new wavelengths on the opposite side of the spectrum. When the process is complete (right), no more temporal overlap between different spectral components is possible, and the broadening stops.

5

CHARACTERIZATION OF SUSPENDED-CORE FIBERS

In the previous chapter the main advantages of ANDi fibers have been presented, arguing that their pulse-preserving, spectrally flat, coherent supercontinuum generation capabilities has opened the door for SC sources to be employed in ultrafast photonics applications, specifically in those areas in which ultrashort coherent pulses are required, like seeding broadband high-power amplifiers [157], or for time-resolved absorption spectroscopy [158].

While PCFs have generally been the main option for dispersion-engineered highly-nonlinear fibers (HNLFs), several fiber designs have been proposed in the recent years. In fact, for the realization of silica PCFs with an ANDi profile, a large number of closely spaced air holes in the cladding is necessary for reducing confinement losses, which results in difficulty in the manufacturing process [31].

In [159], a new suspended-core fiber (SCF) design consisting in a sub-micrometer core suspended in air, connected to the cladding with thin silica bridges, was proposed. These fibers showed a similar level of design freedom compared to PCFs, with their dispersion being controlled through the core diameter and the number of bridges, but proved to be much easier to fabricate. So far they have been tested for SC ANDi generation in the visible range, with MDWs in the vicinity of 500 nm, but with difficulty in tuning it to the near infrared region, where the difficulty in reducing the confinement losses of silica PCFs is higher [31].

In the recent years, following the idea of new microstructured ZBLAN glass fibers [160], a new degree of freedom has been added to the SCF design, consisting in adding a nano-hole in the center of the suspended silica core [161].

With this additional design parameter, the tunability of the MDW was greatly improved, showing the possibility of extending the ANDi regime beyond 2 µm. A comparison with a typical ANDi PCFs geometry is also given in [161], showing that their guiding capability is greatly reduced when the MDW exceeds 1 µm limit. At the same time, SCFs are able to maintain low confinement losses and nonlinear coefficient with over an order of magnitude of improvement compared to similar PCFs.

By elongating one side of the core and using an elliptical nano-hole, extremely birefringent fibers with vastly different dispersion charac-

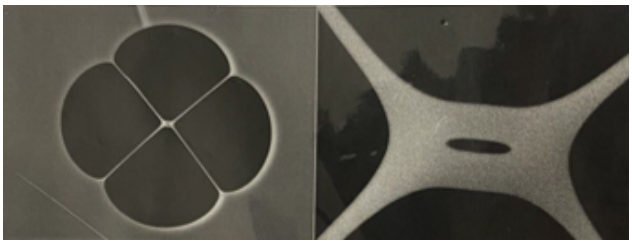


Figure 5.1: Scanning-electron microscope images of the fiber under test, showing the structure of the cladding with silica bridges supporting the core (left) and closeup of the suspended core (right).

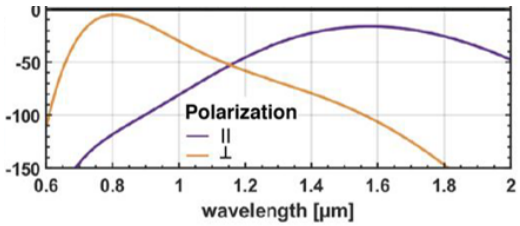


Figure 5.2: Dispersion profiles of the considered SCF for input polarizations aligned (∥) or orthogonal (⊥) to the long core axis.

teristics on the two principal polarization modes can be realized. In principle, these fibers could open the possibility of ultra-broadband ANDi SC using multiple pumps located at different central wavelengths.

In particular, a sample fabricated using 4 silica bridges, a $1.06 \times 2.18\,\mu m$ core, with a $0.16 \times 0.8\,\mu m$ nano-hole, whose microscope images are reported in fig. 5.1, showed ANDi profiles with MDWs located near 1600 nm and 800 nm for polarization modes aligned with the long and short core axis, respectively (fig. 5.2).

In this chapter, preliminary results regarding the experimental characterization of this fiber sample are presented for the first time, showing the difference in terms of generated SC and noise characteristics when pumped near the two MDWs. Fabrication of the fiber sample and the finite-element method (FEM) simulations were performed by the Leibniz Institute of Photonic Technology, Jena, Germany.

5.1 FIBER DISPERSION AND MULTIMODAL CHARACTERISTICS

Although the fiber sample under consideration is designed to support ANDi profiles on the two fundamental mode polarizations, the design parameters used for the fabrication also dictate the presence of guided higher order modes [161].

The dynamics of multimodal SC generation have been a topic of research in the past years, with results in the literature showing a

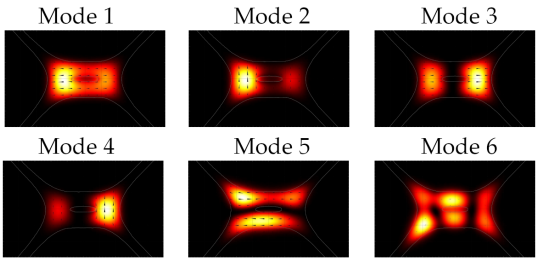


Figure 5.3: Mode intensity distributions and field orientations calculated using FEM simulations at $\lambda = 900$.

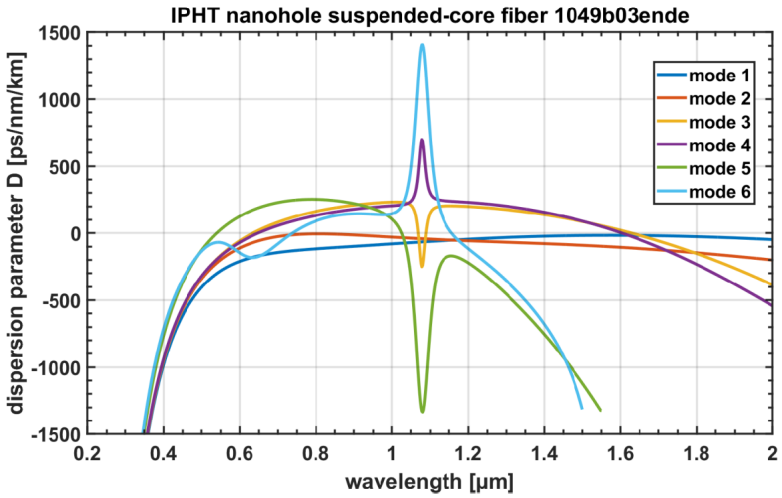


Figure 5.4: Dispersion profile of the first 6 modes calculated using FEM simulations.

significant dependence of the output spectrum on the pumping conditions as a result of mode coupling, intermodal FWM and XPM [162, 163].

Even though the use of multimode fibers for SC generation has enabled unprecedented power handling capabilities and high output beam quality [164, 165], the presence of higher order modes can also pose a limit to the SC bandwidth due to the depletion of the fundamental mode or due to polarization coupling [166, 167].

For the fiber under test, a FEM simulation of the core geometry obtained from scanning-electron microscope images of the suspended-core structure reported in fig. 5.1 revealed the presence of 6 total spatial and polarization modes; their intensity distribution and field orientation at 900 nm is reported in fig. 5.3, while their dispersion and group velocity curves are illustrated in figs. 5.4 and 5.5, respectively.

Apart from the first two ANDi modes (labeled mode 1 and mode 2) with MDWs at 800 nm and 1600 nm, the remaining modes present

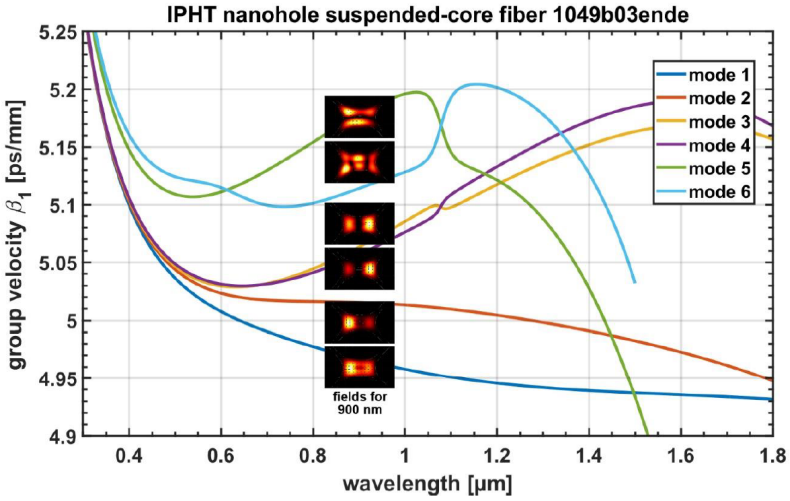


Figure 5.5: Group velocity curve of the first 6 modes calculated using FEM simulations.

dispersion curves with two or more ZDWs, in part caused by the divergence at approximately 1100 nm. This unconventional feature is more accentuated for modes 5 and 6, and contributes to the formation of additional ZDWs at shorter wavelengths than typical dual-ZDW PCF designs, at 1050 nm and 1150 nm respectively. Modes 3 and 4 instead exhibit a less pronounced divergence, with the latter characterized by the second ZDWs at over 1600 nm.

5.2 SC GENERATION USING A TI:SAPPHIRE LASER

As seen in the previous sections, the polarization of the fundamental mode aligned along the short core axis of the SCF presents an ANDi profile with a MDW located near 800 nm. Consequently, the fiber can be efficiently pumped using Titanium Sapphire lasers, whose radiation can be generally tuned from the visible to the near-infrared, covering a a portion of the spectrum from 650 nm to 1100 nm. The specific source used for characterizing the fiber in this spectral region is a Ti:Sapphire mode-locked oscillator (Coherent Chameleon Vision-S) emitting 80 fs pulses with a 80 MHz repetition rate, and with a maximum average power of approximately 4 W at the central wavelength of 800 nm.

The rest of the setup that was built for characterizing the fiber is reported in fig. 5.6.

The femtosecond pulses coming from the pump first pass through an isolator (EURYS Broadband Optical Isolator) in order to avoid unwanted reflections that may cause instability of the mode-locking of the source, and then through a power adjustment stage composed

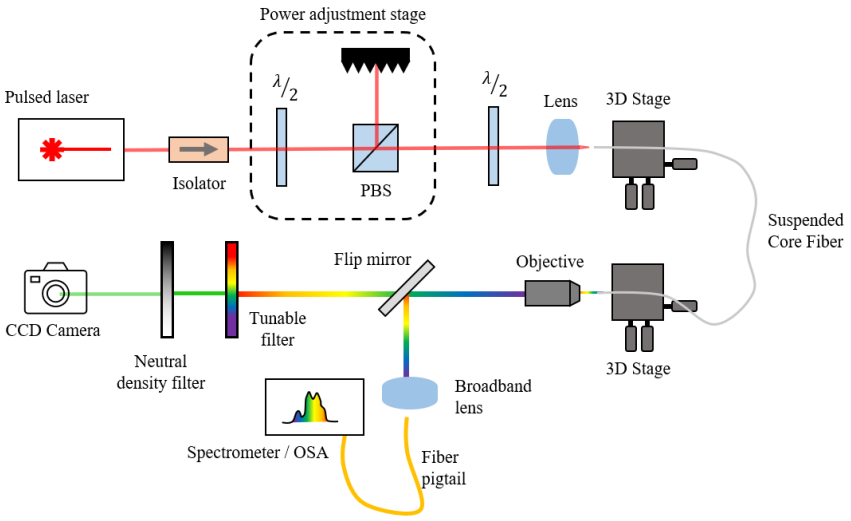


Figure 5.6: Setup for measuring the SC and output modes. Description is given in the text.

by an half-wave plate and a polarization beam splitter (PBS). The attenuated beam is then routed with a series of silver mirrors and periscopes (not illustrated in the figure) to a second half-wave plate for controlling the polarization orientation of light injected in the fiber. Using an aspheric lens, the beam is focused on the core of the SCF, whose start and end are fixed on high-precision 3D mechanical translation stages. The beam coming from the output of the fiber is then collected with a 40X magnification microscope objective and routed to two different paths using a silver flip-mirror. The first path is for the measurement of the generated SC, and consists in a broadband lens which focuses the beam on a large mode area fiber connected either to a spectrometer to measure the portion of spectrum from 300 nm to 900 nm or to an optical spectrum analyzer (OSA) covering the remaining part of the spectrum up to 1700 nm. When acquiring the long-wavelength portion of the spectrum with the OSA, an additional measurement with a long-pass filter with a cutoff wavelength of 1400 nm was used to remove the replicas of the short-wavelength components appearing at twice their wavelength. The three total spectral measurements are then numerically stitched together to obtain the complete SC spectrum.

The second path instead is dedicated to the measurement of the different modes of the SC, and consists in a tunable interference filter to isolate a selected portion of the spectrum, a variable neutral density filter to attenuate the beam, and finally a charged-coupled device (CCD) camera to capture the light distribution. The camera is located several meters from the output fiber to get a sufficiently high beam

Figure 5.7: Fiber input facet after core collapsing and cleaving. Photo taken from the fusion splicer display.

magnification, while the focus adjustment is performed by controlling the distance of the microscope objective from the fiber facet using the output translation stage.

5.2.1 *Coupling light in the SCF*

In order to couple light in the SCF core, the first few centimeters of fiber are first stripped of their protective coating, exposing the cladding, and then hand-cleaved using a ceramic blade. Finally, the fiber is taped to a metal V-groove fiber holder with copper tape to facilitate thermal dissipation and limit the effect of heat on the alignment, and finally fixed to the input translation stage. The same procedure was applied to the fiber end, which is then fixed to the output translation stage using a FC/PC bare fiber adapter.

Due to the small core area of $\approx 2\,\mu m^2$, the coupling efficiency proved to be a limiting factor. Using lenses with focal lengths ranging from 2 mm to 7 mm, only a maximum of 15% of the input power is measured at the fiber output. Additionally, the coupling conditions were extremely sensitive to environmental noise and vibrations, causing rapid changes to the generated SC that prevent its measurement with the procedure previously described.

The coupling efficiency can be greatly increased following the procedure outlined in [168]. The aim is to virtually increase the core of the SCF by splicing it to a "dummy" fiber with a similar cladding diameter, causing a collapse of the suspended-core structure. After cleaving the resulting fiber close to the collapsed core using a precision cleaver, the increased core size of the dummy fiber guides the coupled light in the now tapered SCF.

The collapsed core obtained by splicing the SCF to a 1060 HI single-mode fiber is visualized in fig. 5.7, and it enables coupling efficiencies of almost 60% using an aspheric lens with a 6.1 mm focal length. The stability of the generated SC also increased significantly due to vibrations and environmental changes only marginally affecting the

alignment of the increased core. This procedure is applied only to the input facet of the fiber, while the output facet is simply hand cleaved as initially described in order to preserve the light distributions of the SC modes.

5.2.2 *Evolution of the supercontinuum*

First, the orientation of the fiber axis at the input and output are determined by placing a broadband polarizer at the fiber output and measuring the transmitted power with a power meter. The angle of the input half-wave plate and of the polarizer are iteratively adjusted until the measured power is maximized, corresponding to the case in which light is injected and collected on the same core axis. This procedure is performed by lowering the pump power in order to avoid nonlinear coupling between orthogonal polarizations.

Aligning the half-wave plate with the short core axis, the translation stage was adjusted to maximize spectral broadening. Although the input polarization corresponds to the fundamental ANDi mode with MDW at 800 nm, the resulting SC did not exhibit the typical features associated to the ANDi case for any combination of pump wavelength, input focusing lens, and coupling conditions.

This is illustrated in fig. 5.8, where the spectral broadening dynamics are visualized for different input powers, tuning the Ti:Sapphire laser to 810 nm. The input power was limited to 500 mW, which was determined to be a safe threshold to avoid damaging the collapsed core structure, and sufficiently high to guarantee the completion of the spectral broadening process.

Apart from the initial SPM-induced broadening occurring with limited input power, the evolution of the spectrum is characterized by the presence of several peaks and fine structures, with the first soliton appearing near 600 nm at 12 mW, corresponding to the ZDW of modes 3 and 4 [1].

Increasing the power, more peaks appear in the region from 600 to 900 nm, which transform in a continuum extending up to 1300 nm for 49 mW [2]. On the short-wavelength edge, the formation of prominent peaks can be associated to the generation of dispersive waves, as described in section 4.2.4. Increasing the power leads to the formation of new short-wavelength peaks and their continuous blue-shift associated to the Raman-induced soliton red-shift. The sharp loss peak at

1 The different noise floor levels at short and long wavelengths are the result of measuring the spectrum with multiple instruments.

2 The pedestal appearing at short wavelengths for higher input powers is a measurement artifact of the spectrometer; it has been experimentally verified using an UV filter that no signal was generated in those spectral regions.

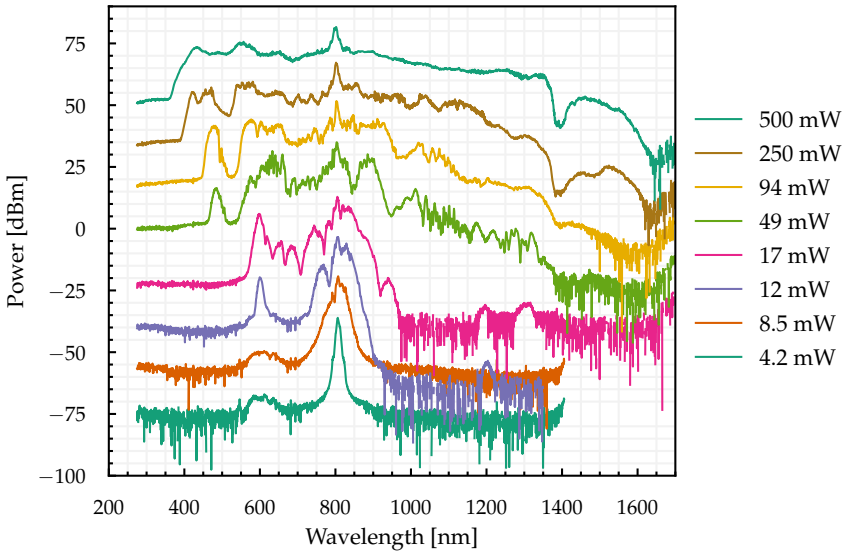


Figure 5.8: Evolution of the output spectrum as a function of the input power, when the fiber is pumped with a Ti:Sapphire laser at 810 nm with input polarization orthogonal to the long core axis.

1400 nm caused by the high OH concentration in silica is also clearly identified.

The broadening process then stops at approximately 1600 nm on the long wavelength edge, where the second ZDW of mode 3 and 4 is located. On the short-wavelength side instead, the SC extends to 350 nm, which corresponds to the shortest wavelength that can be achieved considering the group-velocity matching and dispersive wave trapping arguments detailed in section 4.2.4 and using the curves of fig. 5.5 for modes 3 and 4, resulting in a SC spanning more than two octaves.

The modal content of the SC acquired using the CCD camera reveals the presence of light distributions similar to those of modes 2, 3, and 4 independently of the input power and over most of the spectrum, with intensity patterns covering the two main lobes adjacent to the nano-hole, with the exception of the region near the pump, where the power is higher and the entire suspended-core structure is illuminated, as seen in fig. 5.9. These results suggest preferential coupling of modes with similar intensity distributions, and in absence of ANDi characteristics, support the reasoning for considering the dispersion and group velocity curves of modes 3 and 4 when discussing the broadening dynamics.

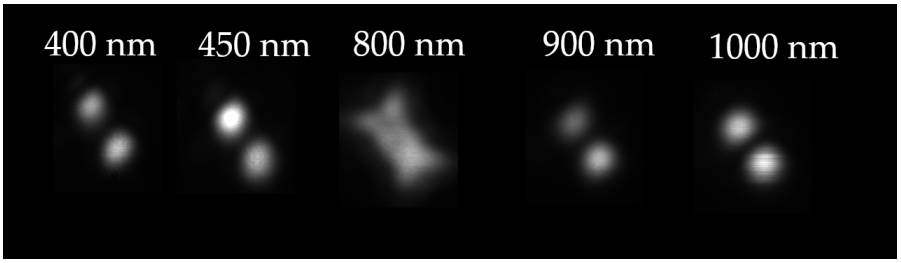


Figure 5.9: Modes measured at the output of the fiber when the input polarization is orthogonal to the long-core axis of the fiber when pumped with a Ti:Sapphire laser at 810 nm, with input power 500 mW.

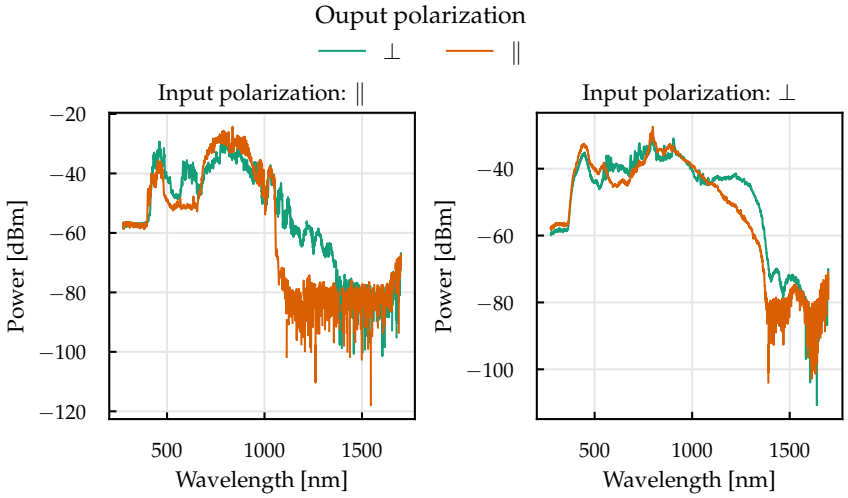


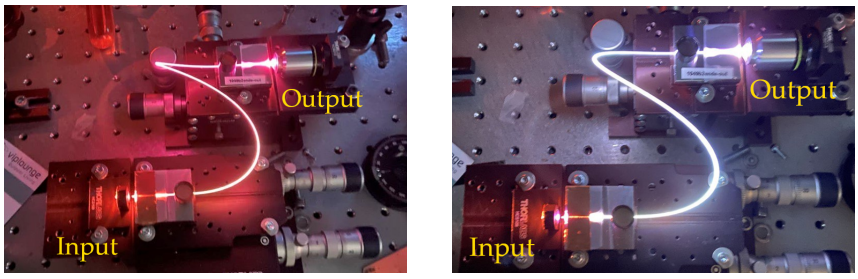
Figure 5.10: Measured polarization extinction ratio when the input polarization is parallel (left) and orthogonal (right) to the long-core axis of the fiber when pumped with a Ti:Sapphire laser at 810 nm with input power 500 mW.

5.2.3 *Polarization extinction ratio*

Using a broadband polarizer at the fiber output, the polarization extinction ratio (PER) can be measured by acquiring the SC on the two fiber axis. These measurements are reported in fig. 5.10 when the input polarization is parallel (left) and orthogonal (right) to the long core axis of the fiber.

In both cases, the PER is close to 0 over the majority of the spectrum, suggesting strong coupling also between modes with orthogonal polarizations. The shape of the generated SC is also appreciably different in the two cases, showing a narrower spectrum both on the short and long wavelength edge and a poorer spectral content in the visible range when the input polarization is aligned with the long core axis.

In this case, the sharp drop-off at approximately 1100 nm is caused by the ZDW of the (5,6) mode pair, which can be more efficiently excited due to their polarization being aligned with the long core axis, and whose presence is detected with the CCD camera.



(a) Input polarization parallel to long core axis.

(b) Input polarization orthogonal to long core axis.

Figure 5.11: Photos of the output SC when the fiber is pumped at 810 nm.

The difference in generated SC when the fiber is pumped on the two axis is also well appreciated by naked eye. In fig. 5.11 are reported the photos of the fiber in the two pumping conditions; when the input polarization is aligned along the short core axis (right), which corresponds to the polarization of the ANDi mode with MDW at 800 nm, the fiber has an almost perfect white glow. On the other hand, when the input polarization is aligned with the long core axis (left), the dip in the green region of the spectrum gives the fiber a pink/purple hue, resulting from the strong red spectral content mixing with the violet/blue peak from 400 to 500 nm.

5.2.4 *RIN measurements*

The coherence degradation of SC sources caused by quantum noise amplification or induced by technical noise of the pump is generally quantified through the computation of the modulus of the complex degree of first-order coherence, defined at each wavelength by

$$|g_{12}^{(1)}(\lambda, t_1 - t_2)| = \left| \frac{\langle E_1^*(\lambda, t_1) E_2(\lambda, t_2) \rangle}{\sqrt{\langle |E_1(\lambda, t_1)|^2 \rangle \langle |E_2(\lambda, t_2)|^2 \rangle}} \right|, \tag{5.1}$$

where the angular brackets denote the ensemble average over independently generated pairs of SC spectra [149]. This quantity is useful to describe the phase stability characteristics of SC spectra and its shot-to-shot fluctuations.

Wavelength-dependent intensity fluctuations can also be quantified through the relative intensity noise (RIN), which is calculated from

the radio-frequency (RF) noise spectrum within a particular optical bandwidth, yielding a value that is directly related to the percentage of pulse-to-pulse amplitude fluctuations [149]. Although $|g_{12}^{(1)}|$ and the RIN describe different things, it has been shown that SC spectra exhibiting strong phase fluctuations also exhibit strong intensity fluctuations, and hence the two parameters are strongly correlated, with the latter being considerable easier to evaluate experimentally [149].

The RIN of the generated SC is thus measured using the setup in fig. 5.6, replacing the CCD camera with a high-speed photodiode, a low-pass filter, and an electronic spectrum analyzer (ESA) to capture the RF noise spectrum. For each noise measurement, the reference DC power was calculated by measuring the root-mean-square (RMS) voltage V_{rms} over a signal trace acquired using a digital oscilloscope as $P_{DC} = V_{rms}^2 / R_L$, where $R_L = 50\,\Omega$ is the load resistance of the ESA. This value is used to normalize the ESA measurements and obtain the final RIN spectrum.

Due to the availability of the equipment, specifically the Si-based detector covering the bandwidth up to 1100 nm and the low-pass filter with a cutoff frequency that is lower than half the repetition rate of the source, the measurements underestimate the true RIN, but can be useful to qualitatively describe the noise-enhancing mechanisms that occur in in the SC generation process.

To this end, the RIN of the total SC is first measured and compared to the source. In fig. 5.12 the RIN spectrum (top) and integrated RIN (bottom) are both reported, with the latter showing an approximately three-fold increase of the SC RIN with respect to the source, approaching 0.65%.

Indeed, from the noise spectrum in fig. 5.12 (top) we can observe an upward shift of the SC trace, attributable to a white noise-like process covering the entire band. This feature is typically observed in conventional SC sources, but can also manifest itself in ANDi fibers due to quantum noise amplification deriving from Raman and parametric FWM gain [31, 157, 169].

Additionally, the significant increase in integrated RIN at higher frequencies can also be explained by noting the more prominent high-frequency components of the SC RIN spectrum, which is generally associated to polarization-related processes such as polarization modulation instability (PMI), causing rapid oscillations in the polarization state of the SC, explaining the poor PER reported in section 5.2.3 [157, 170].

Similar measurements are performed on individual sections of the spectrum by placing the interference filter before the photodiode. The total integrated RIN and the corresponding SC are reported in fig. 5.13.

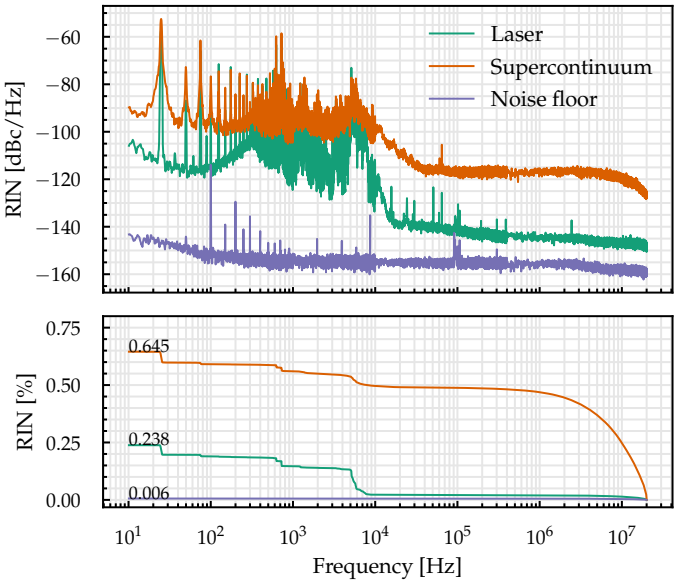


Figure 5.12: RIN spectrum (top) and corresponding integrated RIN (bottom) when the fiber is pumped using a Ti:Sapphire laser at 810 nm with input polarization orthogonal to the long core axis. The measured noise floor includes the contributions of both the photodiode and the ESA.

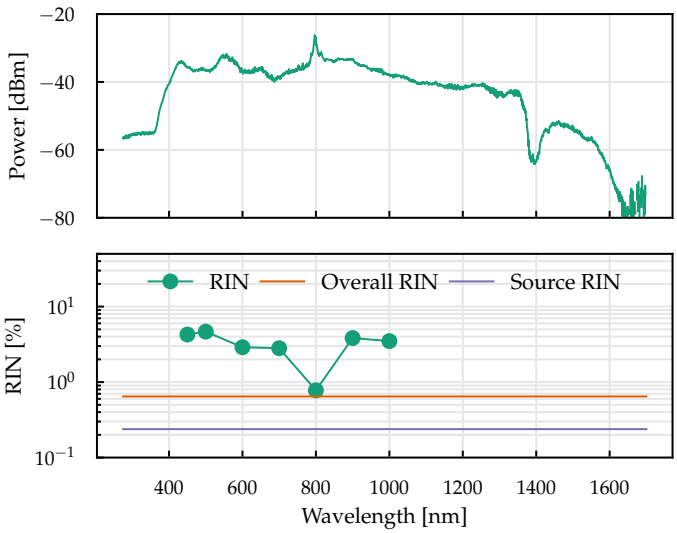


Figure 5.13: Output spectrum (top) and RIN values (bottom) measured at specific wavelengths when the fiber is pumped with a Ti:Sapphire laser at 810 nm, with input polarization orthogonal to the long core axis and input power 290 mW.

Similar to previously reported cases in the literature, both in anomalous and normal dispersion pumping, the total intensity noise is the minimum near the pump wavelength and increases moving towards the edges of the SC, reaching a maximum of nearly 4.7%, corresponding to almost a 20 times increase with respect to the source RIN [171, 172]. The individual RIN spectra highlight similar features to those reported in fig. 5.17 for the SC, with a significant upward shift of the entire spectrum and accentuated high frequency noise.

5.3 SC GENERATION USING AN ERBIUM FIBER LASER

As previously detailed, the fundamental polarization mode aligned along the long core axis of the fiber presents a MDW located at 1600 nm that can be conveniently pumped using Er:fiber lasers, which can emit radiation centered at approximately 1550 nm. The specific source used in this case is characterized by 80 fs pulses at a repetition rate of 40 MHz and average power of 500 mW. The rest of setup described in fig. 5.6 was adapted using equivalent components for this spectral region. In this case, the entire spectrum can be acquired with a single OSA covering the bandwidth from 1200 nm to 2400 nm.

After aligning the input polarization to the long core axis of the fiber using the half-wave plate in order to excite the ANDi mode with MDW at 1600 nm, a series of spectral measurements varying the pump power was acquired.

Contrary to the previous case, here the ANDi regime can be readily achieved; in fact, the characteristic broadening dynamics typical of ANDi SC can be identified from the results reported in fig. 5.14, showing the simple and regular evolution of the spectrum with increasing input power.

In the first stage of SC generation, up to the curve corresponding to input power of 67 mW, the gradual broadening associated with SPM can be observed. With increasing input power, the process of OWB initiates the energy transfer from the center to the edge of the spectrum, as described in section 4.3.1, causing the formation of lower level shoulders that can be detected starting from the curve corresponding to 128 mW.

The reduced power of the laser caused by reflections and losses in the optical setup, the typical higher peak powers required for ANDi SC generation in comparison to conventional SC [31], and the lower nonlinear coefficient of the fiber at longer wavelengths [161], combine in halting the spectral broadening process before it is fully developed, resulting in a SC spanning the region from 1300 nm to approximately 1900 nm.

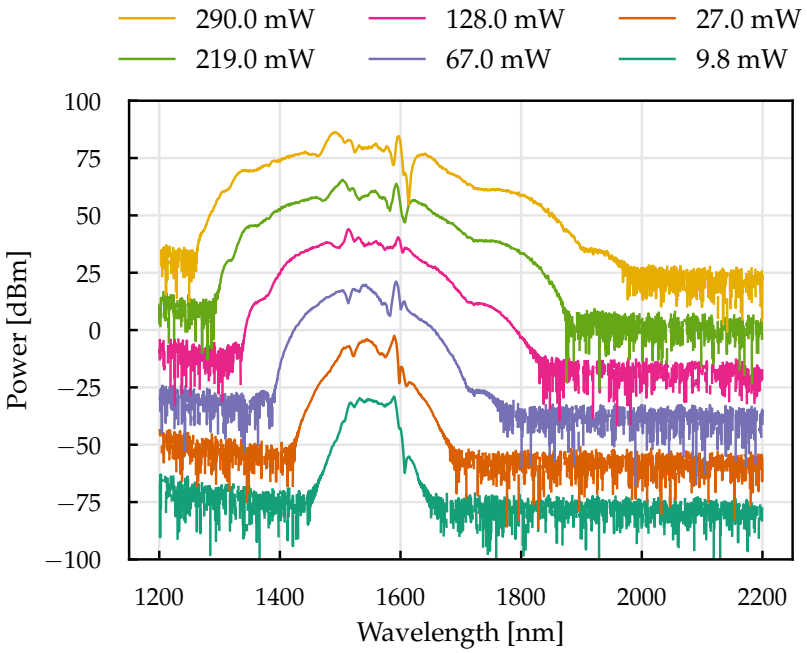


Figure 5.14: Evolution of the SC spectrum varying the input power, when the input polarization is parallel to the long core axis. Pump is an Erbium fiber laser centered at 1550 nm.

5.3.1 *Polarization extinction ratio*

Similarly to previous case, using a broadband polarizer at the output of the SCF the polarization extinction ratio was measured for both input fiber axes, and reported in fig. 5.15.

Compared to the conventional regime reached when pumping the fiber at 800 nm, the achieved ANDi SC shows a significant improvement for both input polarization axes, reaching a PER of nearly 25 dB on the edge of the spectrum when the pump is aligned with the long core axis (indicated with $\parallel$). Instead, when the input polarization is orthogonal to the long core axis (indicated by $\perp$ in the plot), the fiber showed an appreciable reduction of the generated SC spectrum while maintaining its ANDi characteristics; the PER is also decreased compared to the previous case, being less than 10 dB in the central part of the spectrum and approximately 15 dB on the edges.

The ease with which the ANDi was reached is readily explained by inspecting the output modes of the SC, showing the same light distribution over the entire spectrum for both input polarizations. The modes measured at 1550 nm are reported in fig. 5.16, exhibiting a significantly different shape than those measured from 400 nm to 1000 nm

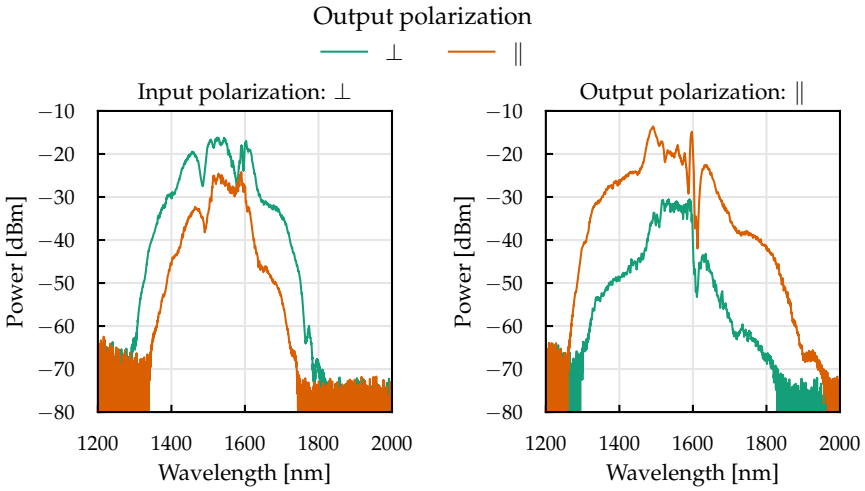
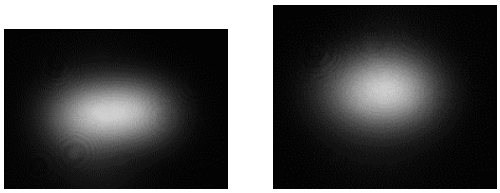


Figure 5.15: Measured polarization extinction ratio when the input polarization is orthogonal (left) and parallel (right) to the long-core axis of the fiber when pumped at 1550 nm with an Er:fiber laser at 290 mW input power.



(a) ⊥ (b) ||

Figure 5.16: Modes captured at the fiber output at 1550 nm for input pump polarization (a) orthogonal and (b) parallel to the long core axis, when pumped with an Er:fiber laser with input power = 290 mW.

when using the Ti:Sapphire laser and from the FEM-computed modes reported in fig. 5.3.

5.3.2 *RIN measurements*

Similarly to the conventional SC regime showed when using the Ti:Sapphire laser, the RIN of the generated SC is measured and compared to that of the source, replacing the photodiode with an InGaAs detector covering the bandwidth from 900 to 2600 nm. The integrated RIN curves graphed in fig. 5.17 (bottom) show that the SC has virtually the same RIN of the Er:fiber laser at about 0.1%, confirming the noise-resilient characteristics of ANDi fibers.

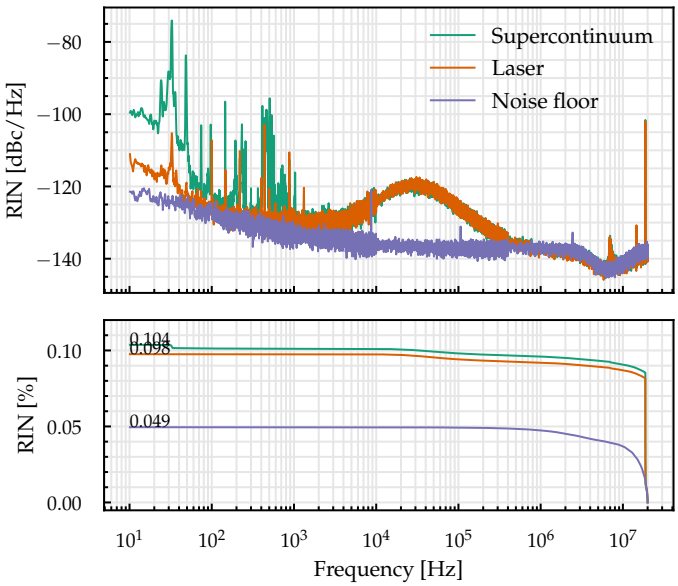


Figure 5.17: RIN spectrum (top) and corresponding integrated RIN (bottom) when the fiber is pumped using a Er:fiber laser at 1550 nmnm.

From the noise spectrum reported in fig. 5.17 (top), no significant quantum noise amplification process is perceived. More specifically, the absence of high-frequency components also suggest a strong suppression of PMI compared to the conventional SC regime, consequence of the excellent PER just discussed. The excess noise peaks that appear in the low frequency range, up to 1 kHz have been observed before with the same source and different fiber samples, and has been attributed to mechanical vibrations at the free-space coupling port of the fiber [173].

Measuring the RIN on different portions of the SC highlights a similar behavior to that encountered in the case of Ti:Sapphire pumping, with fig. 5.18 showing a minimum at the center of the SC, near the pump, and an increase to approximately 5% on the edges of the spectrum, corresponding to an amplification factor of 5X compared to the source RIN. Inspecting the individual noise spectra, the RIN increase is determined by a uniform upward shift of the individual traces with respect to the reference source noise, without additional high-frequency components, suggesting the presence of quantum noise-seeded mixed Raman and parametric gain [169, 170].

5.4 CONCLUSIONS

In this chapter, the preliminary results from the characterization of a SCF sample carried out in collaboration with the University of Bern

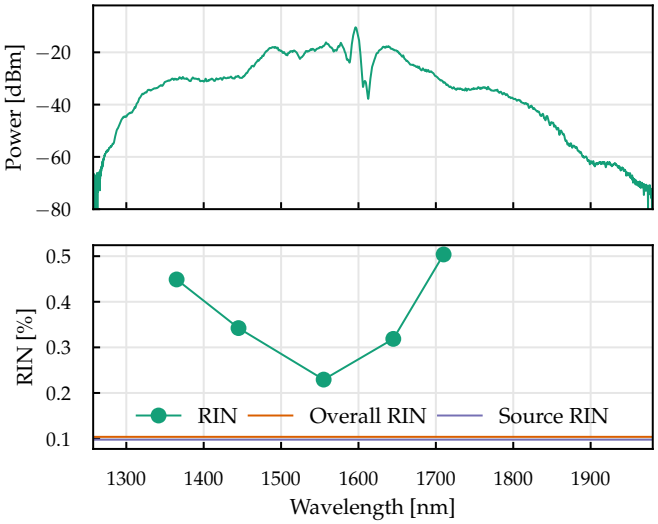


Figure 5.18: Output spectrum (top) and RIN values (bottom) measured at specific wavelengths when pumped with an Erbium fiber laser centered at 1550nm, with input polarization parallel to the long core axis and input power 290 mW.

have been presented. The sample presents ANDi profiles on the two orthogonal polarizations of the fundamental mode with MDWs located at vastly different wavelengths, at 800 nm and 1600 nm respectively. The location of the MDWs is designed to employ widely available Ti:Sapphire and Er:fiber lasers, opening the possibility for the development of ultra-broad dual-pump ANDi SC sources. While the ANDi regime could be readily obtained when the fiber is pumped near its long-wavelength MDW, showing simple spectral broadening dynamics dominated by SPM and OWB, excellent PER, and great noise characteristics, the presence of higher order modes proved to be a limiting factor in stimulating ANDi SC in the visible range using a Ti:Sapphire oscillator. In this case, the generated SC exhibited the typical characteristics of conventional SC, with individual peaks and fine structures appearing in the spectrum as the consequence of the soliton fission process and consequent generation of dispersive waves. Although the final SC spanned two octaves, from 350 nm to approximately 1600 nm, the analysis of its RIN hinted at the presence of polarization-related noise-amplifying processes, which can also be appreciated from the degraded PER on both polarization axes.

6

CONCLUSIONS

In this thesis I presented the main results obtained during my three-year journey as a Ph.D student at the University of Padova, regarding the interplay of nonlinear effects and multimodal propagation in optical fibers.

Thanks to the collaborations with different research groups, I have been able to test my ability on various aspects of fiber optics research, ranging from the application of modern techniques based on machine learning (ML) to analytical and numerical modeling of nonlinear propagation, and concluding with experimental characterization of specialty fibers.

In the context of space-division multiplexing (SDM) communications, different aspects of Raman amplification have been explored.

First, ML techniques have been applied to design and optimization the gain of distributed Raman amplifiers in few-mode fibers (FMFs). Here, neural networks (NNs) have been trained to learn the relationship between a given target gain profile and the corresponding value of pump parameters that are necessary to approximate it. This method has been applied to FMFs supporting up to 6 groups of linearly polarized modes, achieving good results on wide-band wavelength-division multiplexing (WDM) systems in terms of root-mean-square error (RMSE) and gain flatness, minimizing the mode-dependent gain (MDG) of the amplifiers to few percents of the average gain. The robustness of the method has been tested with respect to the change of the input power of the information-bearing signals, showing minimal changes to the predicted gain when signal power fluctuations of 10 dB are present. The noise performance of the amplifiers have also been demonstrated, confirming the superior characteristics of distributed schemes with respect to lumped amplifiers, and showing negligible mode-dependency of its noise figure. The extension to bidirectional pumping schemes has also been tested in order equalize the OSNR tilt that is typical of wide-band counterpumping amplifiers, showing that a good compromise between gain accuracy and OSNR flatness can be achieved by properly designing the cost function that the training algorithms tries to minimize.

The focus has then shifted to the study of the effects of linear coupling caused by stress birefringence and core ellipticity on the gain statistics of few-mode Raman amplifiers. To this aim, the analytical equations that describe the propagation of two different wavelengths

in a few-mode fiber has been derived. In conjunction with the numerical models for birefringence and core ellipticity available in the literature, the equations have been numerically solved to gather a statistical ensemble of the amplifier gain on thousands of fiber realizations. Fibers supporting 2 and 4 groups of linearly polarized modes have been simulated for a wide range of coupling strengths and for two correlation lengths, showing the effects on intra-group and inter-group coupling. Extending the results obtained for the case of polarization mode dispersion (PMD) for single-mode fiber amplifiers, we identify three different coupling regimes. In the low-coupling regime, the average gain of each spatial mode is maximized or minimized depending on the relative orientation of signal and pump polarization, with the gain variance being minimized. As groups start to couple, their average gain converges to a common intermediate value, but with high fluctuations in the output power. When groups are fully coupled, the average gain of the newly formed "supergroup" reaches the minimum value when pump and signal are co-polarized. While the gain variance of the new supergroup is substantially reduced compared to the intermediate coupling regime, the gain variance of the individual modes remains fairly high. The spatial modes of the new supergroup behave similarly to the polarization degeneracies of the fundamental mode of a single-mode fiber: while the total power of the LP_{01} mode is stable at high PMD values, its Stokes vector rapidly moves on the Poincaré sphere, causing the power on the two orthogonal polarizations to fluctuate.

Finally, the last part of my research has been dedicated to the characterization of highly-nonlinear fibers for supercontinuum (SC) generation. In collaboration with the University of Bern, Switzerland, I carried out experimental measurements on a sample of suspended-core fiber manufactured by the Leibniz Institute of Photonic Technology, Jena, Germany, and presented preliminary results involving the measurement of the evolution dynamics of the SC, its polarization-extinction ratio, and its noise characteristics. Although the output SC spanned more than two octaves, the multimodal characteristics proved to be a limiting factor in reaching the low-noise, high-coherence all-normal dispersion (ANDi) regime that these fibers where manufactured for, when pumped using a femtosecond Ti:Sapphire laser at 800 nm. When pumped using an Er:fiber laser at 1550 nm instead, the ANDi SC was generated, showing no impairments related to multimodal interactions, exhibiting an excellent polarization-extinction ratio exceeding 20 dB, and a total relative intensity noise (RIN) virtually equal to that of the source.

6.1 FUTURE WORKS

During these three years, different aspects of nonlinear fiber optics have been investigated and interesting results have been obtained. Nevertheless, there is room for further research, with some details requiring additional work.

Regarding the presented ML method to shape the amplification profile of few-mode fiber Raman amplifiers, when using bidirectional pumps in conjunction with ad-hoc cost functions to flatten the OSNR spectrum, the contribution of nonlinear interference noise (NLIN) on the WDM spectrum could become important. In fact, while amplified spontaneous emission (ASE) noise is generally suppressed when using copropagating pumps, the path-averaged power of the signals is higher when compared to counterpropagating schemes, meaning that cross-phase modulation (XPM) between WDM channels is enhanced. A more complete approach should account for these effects by computing the NLIN using the models presented in [174, 175], and embedding them in the differentiable solver. This would allow the NNs to learn how to optimally distribute the power between forward- and backward-propagating pumps, provided that the cost function used to train them has been appended with an additional term that is proportional to the NLIN.

Moreover, although the study on effects of linear coupling on Raman amplification highlighted some interesting results, the analysis has been limited only to perturbations that are intrinsic to the manufacturing process of optical fibers. Extrinsic effects, such as bending and twisting, could be readily studied. Additionally, the presented results are limited to the case of the Raman pump co-propagating with the transmitted signals; with proper changes to the numerical models used, the case of counterpropagating pumps could be studied.

Finally, regarding the research on supercontinuum generation, while the experimental measurements are necessary for the characterization of the fiber samples, they must be combined with numerical simulations in order to fully understand the interplay between the main nonlinear effects responsible for spectrum broadening and inter-modal propagation. This aspect is subject of ongoing research in collaboration with the University of Bern.

BIBLIOGRAPHY

[1] T. H. Maiman. "Stimulated Optical Radiation in Ruby." In: *Nature* 187.4736 (Aug. 1960), pp. 493–494.

[2] Jeff Hecht. "The Remarkable Fiber Optic Vision Of Charles Kao." In: *Optics & Photonics News* March 2019 (2019), p. 8.

[3] Jeff Hecht. *City of light : the story of fiber optics*. English. Oxford University Press New York, 1999, xii, 316 p. : ISBN: 0195108183.

[4] F. P. Kapron, D. B. Keck, and R. D. Maurer. "Radiation Losses in Glass Optical Waveguides." In: *Applied Physics Letters* 17.10 (Nov. 1970), pp. 423–425.

[5] T. Miya, Y. Terunuma, T. Hosaka, and T. Miyashita. "Ultimate Low-Loss Single-Mode Fibre at 1.55 *M*m." In: *Electronics Letters* 15.4 (Feb. 1979), pp. 106–108.

[6] Mitch Jacoby. *As Telecom Demands Grow, Optical Fibers Will Need to Level Up*. Mar. 2020.

[7] TeleGeography. *Submarine Cable Map, www.submarinecablemap.com*.

[8] D. J. Richardson. "New Optical Fibres for High-Capacity Optical Communications." In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374.2062 (Mar. 2016), p. 20140441.

[9] René-Jean Essiambre, Gerhard Kramer, Peter J. Winzer, Gerard J. Foschini, and Bernhard Goebel. "Capacity Limits of Optical Fiber Networks." In: *Journal of Lightwave Technology* 28.4 (Feb. 2010), pp. 662–701.

[10] D. J. Richardson, J. M. Fini, and L. E. Nelson. "Space-Division Multiplexing in Optical Fibres." In: *Nature Photonics* 7.5 (May 2013), pp. 354–362.

[11] Benjamin J. Puttnam, Ruben S. Luís, Georg Rademacher, Manuel Mendez-Astudilio, Yoshinari Awaji, and Hideaki Furukawa. "S, C and Extended L-Band Transmission with Doped Fiber and Distributed Raman Amplification." In: *2021 Optical Fiber Communications Conference and Exhibition (OFC)*. June 2021, pp. 1–3.

[12] A. D. Ellis, N. Mac Suibhne, D. Saad, and D. N. Payne. "Communication Networks beyond the Capacity Crunch." In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374.2062 (Mar. 2016), p. 20150191.

[13] Inc Cisco Systems. *Cisco Visual Networking Index: Forecast and Trends, 2017–2022*. 2019.

[14] Andrew Chralyvy. "The Coming Capacity Crunch." In: *European Conference on Optical Communication (ECOC)*. Vienna, Austria, Sep. 2009, Plenary talk.

[15] L. Palmieri and A. Galtarossa. "Coupling Effects Among Degenerate Modes in Multimode Optical Fibers." In: *IEEE Photonics Journal* 6.6 (Dec. 2014), pp. 1–8.

[16] R.-J. Essiambre, R. Ryf, N. K. Fontaine, and S. Randel. "Breakthroughs in Photonics 2012: Space-Division Multiplexing in Multimode and Multicore Fibers for High-Capacity Optical Communication." In: *IEEE Photonics Journal* 5.2 (Apr. 2013), pp. 0701307–0701307.

[17] T. Torounidis, P.A. Andrekson, and B.-E. Olsson. "Fiber-Optical Parametric Amplifier with 70-dB Gain." In: *IEEE Photonics Technology Letters* 18.10 (May 2006), pp. 1194–1196.

[18] V. Sasikala and K. Chitra. "All Optical Switching and Associated Technologies: A Review." In: *Journal of Optics* 47.3 (Sept. 2018), pp. 307–317.

[19] H. J Eichler, A Mocofanescu, Th Riesbeck, E Risse, and D Bedau. "Stimulated Brillouin Scattering in Multimode Fibers for Optical Phase Conjugation." In: *Optics Communications* 208.4 (July 2002), pp. 427–431.

[20] Mohammed N. Islam, ed. *Raman Amplifiers for Telecommunications 1: Physical Principles*. Springer Series in Optical Sciences 90/1. New York: Springer, 2004.

[21] J. Bromage. "Raman Amplification for Fiber Communications Systems." In: *Journal of Lightwave Technology* 22.1 (Jan. 2004), pp. 79–93.

[22] K. Suzuki, H. Ono, T. Mizuno, Y. Hashizume, Y. Abe, T. Takahashi, K. Takenaga, S. Matsuo, and H. Takara. "Pump Light Source for Distributed Raman Amplification in MCFs With LD Sharing Circuit." In: *IEEE Photonics Technology Letters* 24.21 (Nov. 2012), pp. 1937–1940.

[23] Govind P. Agrawal. "Nonlinear Fiber Optics: Its History and Recent Progress [Invited]." In: *Journal of the Optical Society of America B* 28.12 (Dec. 2011), A1.

[24] Roger H. Stolen. "The Early Years of Fiber Nonlinear Optics." In: *Journal of Lightwave Technology* 26.9 (May 2008), pp. 1021–1031.

[25] Philip St.J. Russell. "Photonic-Crystal Fibers." In: *Journal of Lightwave Technology* 24.12 (Dec. 2006), pp. 4729–4749.

[26] Robert R. Alfano. "The Supercontinuum Laser Source: The Ultimate White Light." In: *Scientific American* 295.6 (Dec. 2006), pp. 86–93.

[27] Sucbei Moon and Dug Young Kim. "Ultra-High-Speed Optical Coherence Tomography with a Stretched Pulse Supercontinuum Source." In: *Optics Express* 14.24 (2006), p. 11575.

[28] John T. Woodward, Allan W. Smith, Colleen A. Jenkins, Chungsan Lin, Steven W. Brown, and Keith R. Lykke. "Supercontinuum Sources for Metrology." In: *Metrologia* 46.4 (June 2009), S277–S282.

[29] S. A. Kovalenko, A. L. Dobryakov, J. Ruthmann, and N. P. Ernsting. "Femtosecond Spectroscopy of Condensed Phases with Chirped Supercontinuum Probing." In: *Physical Review A* 59.3 (Mar. 1999), pp. 2369–2384.

[30] Takuya Ohara, Hidehiko Takara, Takashi Yamamoto, Hiroji Masuda, Toshio Morioka, Makoto Abe, and Hiroshi Takahashi. "Over-1000-Channel Ultradense WDM Transmission With Supercontinuum Multicarrier Source." In: *Journal of Lightwave Technology* 24.6 (June 2006), p. 2311.

[31] Alexander M Heidt, Dirk-Mathys Spangenberg, Alexander Hartung, and Hartmut Bartelt. "All-Normal Dispersion Fiber Supercontinuum – Principles, Design, and Applications of a Unique White Light Source." In: *The Supercontinuum Laser Source: The Ultimate White Light*, p. 38.

[32] René-Jean Essiambre, Gerard J. Foschini, Gerhard Kramer, and Peter J. Winzer. "Capacity Limits of Information Transport in Fiber-Optic Networks." In: *Physical Review Letters* 101.16 (Oct. 2008), p. 163901.

[33] Roland Ryf et al. "Mode-Division Multiplexing Over 96 Km of Few-Mode Fiber Using Coherent 6×6 MIMO Processing." In: *Journal of Lightwave Technology* 30.4 (Feb. 2012), pp. 521–531.

[34] V.A.J.M. Sleiffer et al. "737 Tb/s (96 x 3 x 256-Gb/s) Mode-Division-Multiplexed DP-16QAM Transmission with Inline MM-EDFA." In: *Optics Express* 20.26 (Dec. 2012), B428.

[35] Neng Bai et al. "Mode-Division Multiplexed Transmission with Inline Few-Mode Fiber Amplifier." In: *Optics Express* 20.3 (Jan. 2012), p. 2668.

[36] R. Ryf et al. "Mode-Equalized Distributed Raman Amplification in 137-Km Few-Mode Fiber." In: *European Conference on Optical Communication (ECOC)*. Geneva, Switzerland, Sep. 2011, Paper Th.13.K.5.

[37] Mina Esmaeelpour, Roland Ryf, Nicolas K. Fontaine, Haoshuo Chen, Alan H. Gnauck, Rene-Jean Essiambre, Jean Toulouse, Yi Sun, and Robert Lingle. "Transmission Over 1050-Km Few-Mode Fiber Based on Bidirectional Distributed Raman Amplification." In: *Journal of Lightwave Technology* 34.8 (Apr. 2016), pp. 1864–1871.

[38] Jiaxiong Li et al. "Experimental Demonstration of a Few-Mode Raman Amplifier with a Flat Gain Covering 1530–1605 Nm." In: *Optics Letters* 43.18 (Sept. 2018), p. 4530.

[39] Jiaxiong Li et al. "Ultra-Low-Noise Mode-Division Multiplexed WDM Transmission Over 100-Km FMF Based on a Second-Order Few-Mode Raman Amplifier." In: *Journal of Lightwave Technology* 36.16 (Aug. 2018), pp. 3254–3260.

[40] Jiaxiong Li, Jiangbing Du, Lin Ma, Ming-Jun Li, Ke Xu, and Zuyuan He. "Second-Order Few-Mode Raman Amplifier for Mode-Division Multiplexed Optical Communication Systems." In: *Optics Express* 25.2 (Jan. 2017), p. 810.

[41] Dagong Jia, Haiwei Zhang, Zhe Ji, Neng Bai, and Guifang Li. "Optical Fiber Amplifiers for Space-Division Multiplexing." In: *Frontiers of Optoelectronics* 5.4 (Dec. 2012), pp. 351–357.

[42] Clifford Headley and Govind P. Agrawal. *Raman Amplification in Fiber Optical Communication Systems*. Academic Press, 2005.

[43] V.E. Perlin and H.G. Winful. "Optimal Design of Flat-Gain Wide-Band Fiber Raman Amplifiers." In: *Journal of Lightwave Technology* 20.2 (Feb./2002), pp. 250–254.

[44] Hai-ming Jiang, Kang Xie, and Ya-fei Wang. "Pump Scheme for Gain-Flattened Raman Fiber Amplifiers Using Improved Particle Swarm Optimization and Modified Shooting Algorithm." In: *Optics Express* 18.11 (May 2010), p. 11033.

[45] B. Neto, A. L. Teixeira, N. Wada, and P. S. André. "Efficient Use of Hybrid Genetic Algorithms in the Gain Optimization of Distributed Raman Amplifiers." In: *Optics Express* 15.26 (2007), p. 17520.

[46] Junhe Zhou. "An Analytical Approach for Gain Optimization in Multimode Fiber Raman Amplifiers." In: *Optics Express* 22.18 (Sept. 2014), p. 21393.

[47] Darko Zibar, Ann Margareth Rosa Brusin, Uiara C. de Moura, Francesco Da Ros, Vittorio Curri, and Andrea Carena. "Inverse System Design Using Machine Learning: The Raman Amplifier Case." In: *Journal of Lightwave Technology* 38.4 (Feb. 2020), pp. 736–753.

[48] U. C. de Moura, F. Da Ros, D. Zibar, A. M. Rosa Brusin, and A. Carena. "Optimization of Raman Amplifiers Using Machine Learning." In: *2021 IEEE Photonics Society Summer Topicals Meeting Series (SUM)*. July 2021, pp. 1–2.

[49] F. Da Ros, U. C. de Moura, R. S. Luis, G. Rademacher, B. J. Puttnam, A. M. Rosa Brusin, A. Carena, Y. Awaji, H. Furukawa, and D. Zibar. "Optimization of a Hybrid EDFA-Raman C+L Band Amplifier through Neural-Network Models." In: *Optical Fiber Communication Conference (OFC) 2021 (2021), Paper Tu1E.5*. Optical Society of America, June 2021, Tu1E.5.

[50] Uiara C. de Moura, Francesco Da Ros, A. Margareth Rosa Brusin, Andrea Carena, and Darko Zibar. "Experimental Characterization of Raman Amplifier Optimization Through Inverse System Design." In: *Journal of Lightwave Technology* 39.4 (Feb. 2021), pp. 1162–1170.

[51] Uiara Celine de Moura, Ann Margareth Rosa Brusin, Andrea Carena, Darko Zibar, and Francesco Da Ros. "Simultaneous Gain Profile Design and Noise Figure Prediction for Raman Amplifiers Using Machine Learning." In: *Optics Letters* 46.5 (Mar. 2021), pp. 1157–1160.

[52] Mehran Soltani, Francesco Da Ros, Andrea Carena, and Darko Zibar. "Inverse Design of a Raman Amplifier in Frequency and Distance Domains Using Convolutional Neural Networks." In: *Optics Letters* 46.11 (June 2021), p. 2650.

[53] Yufeng Chen, Jiangbing Du, Yuting Huang, Ke Xu, and Zuyuan He. "Intelligent Gain Flattening in Wavelength and Space Domain for FMF Raman Amplification by Machine Learning Based Inverse Design." In: *Optics Express* 28.8 (Apr. 2020), pp. 11911–11920.

[54] Gianluca Marcon, Andrea Galtarossa, Luca Palmieri, and Marco Santagiustina. "Model-aware deep learning method for Raman amplification in few-mode fibers." In: (unpublished).

[55] Gianluca Marcon, Andrea Galtarossa, Luca Palmieri, and Marco Santagiustina. "Gain design of few-mode fiber Raman amplifiers using an autoencoder-based machine learning approach." In: *European Conference on Optical Communication (ECOC)*. Brussels, Belgium, 2020, in press.

[56] Roland Ryf, Rene Essiambre, Johannes von Hoyningen-Huene, and Peter Winzer. "Analysis of Mode-Dependent Gain in Raman Amplified Few-Mode Fiber." In: *Proceedings of Optical Fiber Communication Conference (OFC)*. Los Angeles, CA, USA, 2012.

[57] Cristian Antonelli, Antonio Mecozzi, and Mark Shtaif. "Raman Amplification in Multimode Fibers with Random Mode Coupling." In: *Optics Letters* 38.8 (Apr. 2013), p. 1188.

[58] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[59] Dawn Hollenbeck and Cyrus D. Cantrell. "Multiple-Vibrational-Mode Model for Fiber-Optic Raman Gain Spectrum and Response Function." In: *Journal of the Optical Society of America B* 19.12 (Dec. 2002), p. 2886.

[60] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. Cambridge, Massachusetts: The MIT Press, Nov. 2016.

[61] Michael Bartholomew-Biggs, Steven Brown, Bruce Christianson, and Laurence Dixon. "Automatic Differentiation of Algorithms." In: *Journal of Computational and Applied Mathematics*. Numerical Analysis 2000. Vol. IV: Optimization and Nonlinear Equations 124.1 (Dec. 2000), pp. 171–190.

[62] M. Raissi, P. Perdikaris, and G.E. Karniadakis. "Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations." In: *Journal of Computational Physics* 378 (Feb. 2019), pp. 686–707.

[63] Adam Paszke et al. "PyTorch: An Imperative Style, High-Performance Deep Learning Library." In: *Advances in Neural Information Processing Systems 32*. Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. dAlché-Buc, E. Fox, and R. Garnett. Curran Associates, Inc., 2019, pp. 8024–8035.

[64] Atilim Gunes Baydin, Barak A. Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind. "Automatic Differentiation in Machine Learning: A Survey." In: *Journal of Machine Learning Research* 18.153 (2018), pp. 1–43.

[65] Boris Karanov, Mathieu Chagnon, Felix Thouin, Tobias A. Eriksson, Henning Bulow, Domanic Lavery, Polina Bayvel, and Laurent Schmalen. "End-to-End Deep Learning of Optical Fiber Communications." In: *Journal of Lightwave Technology* 36.20 (Oct. 2018), pp. 4843–4855.

[66] Rasmus T. Jones, Tobias A. Eriksson, Metodi P. Yankov, and Darko Zibar. "Deep Learning of Geometric Constellation Shaping Including Fiber Nonlinearities." In: *Proceedings of European Conference on Optical Communication (ECOC)*. Rome, Italy, Sep. 2018, Paper We1F.5.

[67] Ognjen Jovanovic, Metodi P. Yankov, Francesco Da Ros, and Darko Zibar. "End-to-End Learning of a Constellation Shape Robust to Variations in SNR and Laser Linewidth." In: *arXiv:2106.00431 [eess]* (June 2021). arXiv: 2106.00431 [eess].

[68] Simone Gaiarin, Francesco Da Ros, Rasmus T. Jones, and Darko Zibar. "End-to-End Optimization of Coherent Optical Communications Over the Split-Step Fourier Method Guided by the Nonlinear Fourier Transform Theory." In: *Journal of Lightwave Technology* 39.2 (Jan. 2021), pp. 418–428.

[69] Vladislav Neskorniuk, Andrea Carnio, Vinod Bajaj, Domenico Marsella, Sergei K. Turitsyn, Jaroslaw E. Prilepsky, and Vahid Aref. "End-to-End Deep Learning of Long-Haul Coherent Optical Fiber Communications via Regular Perturbation Model." In: *2021 European Conference on Optical Communication (ECOC)*. Sept. 2021, pp. 1–4.

[70] Junho Cho and Peter J. Winzer. "Probabilistic Constellation Shaping for Optical Fiber Communications." In: *Journal of Lightwave Technology* 37.6 (Mar. 2019), pp. 1590–1607.

[71] Diederik P. Kingma and Jimmy Ba. "Adam: A Method for Stochastic Optimization." In: *arXiv:1412.6980 [cs]* (Jan. 2017). arXiv: 1412.6980 [cs].

[72] Xavier Glorot and Yoshua Bengio. "Understanding the Difficulty of Training Deep Feedforward Neural Networks." In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. pp. 249–256, Mar. 2010. Chap. Machine Learning.

[73] Hai Ming Jiang and Kang Xie. "Efficient and Robust Shooting Algorithm for Numerical Design of Bidirectionally Pumped Raman Fiber Amplifiers." In: *Journal of the Optical Society of America B* 29.1 (Jan. 2012), p. 8.

[74] T. Muciaccia, F. Gargano, and V. M. N. Passaro. "A Reconfigurable Optical Metro-Access Network and an Innovative ROADM for Efficient Dynamical Bandwidth Allocation." In: (Jan. 2015), pp. 4 .–4 .

[75] Chunlei Sun, Yu Yu, Guanyu Chen, and Xinliang Zhang. "On-Chip Switch for Reconfigurable Mode-Multiplexing Optical Network." In: *Optics Express* 24.19 (Sept. 2016), pp. 21722–21728.

[76] Yingjie Liu et al. "Arbitrarily Routed Mode-Division Multiplexed Photonic Circuits for Dense Integration." In: *Nature Communications* 10.1 (July 2019), p. 3263.

[77] Wei Wang, Jian Zhao, Zhiqun Yang, Chao Li, Zhen Wang, Liyao Yu, and Ruilong Mi. "Amplified Spontaneous Emission and Rayleigh Scattering in Few-Mode Fiber Raman Amplifiers." In: *IEEE Photonics Technology Letters* 29.14 (July 2017), pp. 1159–1162.

[78] B. Zhu et al. "3.2Tb/s (80 × 42.7Gb/s) Transmission over 20 × 100km of Non-Zero Dispersion Fiber with Simultaneous C + L-band Dispersion Compensation." In: *Optical Fiber Communications Conference (2002), Paper FC8*. Optical Society of America, Mar. 2002, FC8.

[79] S. Kado, Y. Emori, S. Namiki, N. Tsukiji, J. Yoshida, and T. Kimura. "Broadband Flat-Noise Raman Amplifier Using Low-Noise Bidirectionally Pumping Sources." In: *Proceedings 27th European Conference on Optical Communication (Cat. No.01TH8551)*. Vol. 6. Sept. 2001, 38–39 vol.6.

[80] Filipe Ferreira, Stylianos Sygletos, and Andrew Ellis. "Impact of Linear Mode Coupling on the Group Delay Spread in Few-Mode Fibers." In: *Optical Fiber Communication Conference (2015), Paper Tu2D.1*. Optical Society of America, Mar. 2015, Tu2D.1.

[81] Cristian Antonelli, Antonio Mecozzi, Mark Shtaif, and Peter J. Winzer. "Random Coupling between Groups of Degenerate Fiber Modes in Mode Multiplexed Transmission." In: *Optics Express* 21.8 (Apr. 2013), p. 9484.

[82] Gianluca Guerra, Matteo Lonardi, Andrea Galtarossa, Leslie A. Rusch, Alberto Bononi, and Luca Palmieri. "Analysis of Modal Coupling Due to Birefringence and Ellipticity in Strongly Guiding Ring-Core OAM Fibers." In: *Optics Express* 27.6 (Mar. 2019), p. 8308.

[83] Cristian Antonelli, Tetsuya Hayashi, and Antonio Mecozzi. "Random Polarization-Mode Coupling Explains Inter-Core Crosstalk in Uncoupled Multi-Core Fibers." In: *2020 European Conference on Optical Communications (ECOC)*. Dec. 2020, pp. 1–4.

[84] Sercan Ö. Arık and Joseph M. Kahn. "Coupled-Core Multi-Core Fibers for Spatial Multiplexing." In: *IEEE Photonics Technology Letters* 25.21 (Nov. 2013), pp. 2054–2057.

[85] Tetsuya Hayashi, Roland Ryf, Nicolas K. Fontaine, Cen Xia, Sebastian Randel, René-Jean Essiambre, Peter J. Winzer, and Takashi Sasaki. "Coupled-Core Multi-Core Fibers: High-spatial-density Optical Transmission Fibers with Low Differential Modal Properties." In: *2015 European Conference on Optical Communication (ECOC)*. Sept. 2015, pp. 1–3.

[86] Roland Ryf et al. "Long-Distance Transmission over Coupled-Core Multicore Fiber." In: *ECOC 2016 - Post Deadline Paper; 42nd European Conference on Optical Communication*. Sept. 2016, pp. 1–3.

[87] Huiyuan Liu, He Wen, He Wen, Rodrigo Amezcua-Correa, Pierre Sillard, Guifang Li, and Guifang Li. "Reducing Group Delay Spread in a 9-LP Mode FMF Using Uniform Long-Period Gratings." In: *Optical Fiber Communication Conference (2017), Paper Tu2J.5*. Optical Society of America, Mar. 2017, Tu2J.5.

[88] Keang-Po Ho and Joseph M. Kahn. "Linear Propagation Effects in Mode-Division Multiplexing Systems." In: *Journal of Lightwave Technology* 32.4 (Feb. 2014), pp. 614–628.

[89] R. Ryf et al. "Space-Division Multiplexed Transmission over 4200-Km 3-Core Microstructured Fiber." In: *Optical Fiber Communication Conference (2012), Paper PDP5C.2*. Optical Society of America, Mar. 2012, PDP5C.2.

[90] A. Turukhin et al. "Demonstration of 0.52 Pb/s Potential Transmission Capacity over 8,830 Km Using Multicore Fiber." In: *ECOC 2016; 42nd European Conference on Optical Communication*. Sept. 2016, pp. 1–3.

[91] Roland Ryf et al. "Long-Haul Transmission over Multi-Core Fibers with Coupled Cores." In: *2017 European Conference on Optical Communication (ECOC)*. Sept. 2017, pp. 1–3.

[92] Georg Rademacher et al. "High Capacity Transmission in a Coupled-Core Three-Core Multi-Core Fiber." In: *Journal of Lightwave Technology* 39.3 (Feb. 2021), pp. 757–762.

[93] Zhishen Zhang, Jiulin Gan, Xiaobo Heng, Muqiao Li, Jiong Li, Shanhui Xu, and Zhongmin Yang. "Low-Crosstalk Orbital Angular Momentum Fiber Coupler Design." In: *Optics Express* 25.10 (May 2017), p. 11200.

[94] Jingxing Zhang et al. "SDM Transmission of Orbital Angular Momentum Mode Channels over a Multi-Ring-Core Fibre." In: *Nanophotonics* (Nov. 2021).

[95] Katsunori Imamura, Harumi Inaba, Kazunori Mukasa, and Ryuichi Sugizaki. "Weakly Coupled Multi Core Fibers with Optimum Design to Realize Selectively Ecitation of Individual Super-modes." In: *Optical Fiber Communication Conference (2012), Paper OM2D.6*. Optical Society of America, Mar. 2012, OM2D.6.

[96] Taiji Sakamoto, Takayoshi Mori, Masaki Wada, Takashi Yamamoto, Takashi Matsui, Kazuhide Nakajima, and Fumihiko Yamamoto. "Experimental and Numerical Evaluation of Inter-Core Differential Mode Delay Characteristic of Weakly-Coupled Multi-Core Fiber." In: *Optics Express* 22.26 (Dec. 2014), pp. 31966–31976.

[97] Bruno R. P. Pinheiro, João L. Rebola, and Adolfo V. T. Cartaxo. "Analysis of Inter-Core Crosstalk in Weakly-Coupled Multi-Core Fiber Coherent Systems." In: *Journal of Lightwave Technology* 39.1 (Jan. 2021), pp. 42–54.

[98] Biswanath Mukherjee, Ioannis Tomkos, Massimo Tornatore, Peter Winzer, and Yongli Zhao, eds. *Springer Handbook of Optical Networks*. Springer Handbooks. Cham: Springer International Publishing, 2020.

[99] Francesco Poletti and Peter Horak. "Description of Ultrashort Pulse Propagation in Multimode Optical Fibers." In: *Journal of the Optical Society of America B* 25.10 (Oct. 2008), p. 1645.

[100] M. Kolesik and J. V. Moloney. "Nonlinear Optical Pulse Propagation Simulation: From Maxwell's to Unidirectional Equations." In: *Physical Review E* 70.3 (Sept. 2004), p. 036604.

[101] Antonio Mecozzi, Cristian Antonelli, and Mark Shtaif. "Nonlinear Propagation in Multi-Mode Fibers in the Strong Coupling Regime." In: *Optics Express* 20.11 (May 2012), p. 11673.

[102] D. Marcuse, C.R. Manyuk, and P.K.A. Wai. "Application of the Manakov-PMD Equation to Studies of Signal Propagation in Optical Fibers with Randomly Varying Birefringence." In: *Journal of Lightwave Technology* 15.9 (Sept. 1997), pp. 1735–1746.

[103] Sami Mumtaz, René-Jean Essiambre, and Govind P. Agrawal. "Nonlinear Propagation in Multimode and Multicore Fibers: Generalization of the Manakov Equations." In: *Journal of Lightwave Technology* 31.3 (Feb. 2013), pp. 398–406.

[104] Antonio Mecozzi, Cristian Antonelli, and Mark Shtaif. "Coupled Manakov Equations in Multimode Fibers with Strongly Coupled Groups of Modes." In: *Optics Express* 20.21 (Oct. 2012), pp. 23436–23441.

[105] S. Mumtaz, R. Essiambre, and G. P. Agrawal. "Reduction of Nonlinear Penalties Due to Linear Coupling in Multicore Optical Fibers." In: *IEEE Photonics Technology Letters* 24.18 (Sept. 2012), pp. 1574–1576.

[106] C. Antonelli, A. Mecozzi, and M. Shtaif. "Scaling of Inter-Channel Nonlinear Interference Noise and Capacity with the Number of Strongly Coupled Modes in SDM Systems." In: *Optical Fiber Communication Conference*. Anaheim, California: OSA, 2016, W4I.2.

[107] Yuzhe Xiao, René-Jean Essiambre, Marc Desgroseilliers, Antonia M. Tulino, Roland Ryf, Sami Mumtaz, and Govind P. Agrawal. "Theory of Intermodal Four-Wave Mixing with Random Linear Mode Coupling in Few-Mode Fibers." In: *Optics Express* 22.26 (Dec. 2014), p. 32039.

[108] D. E. Ceballos-Herrera, R. Gutierrez-Castrejon, and J. A. Alvarez-Chavez. "Stimulated Raman Scattering and Four-Wave Mixing Effects on Crosstalk of Multicore Fibers." In: *IEEE Photonics Technology Letters* 30.1 (Jan. 2018), pp. 63–66.

[109] Luca Palmieri. "Coupling Mechanism in Multimode Fibers." In: *Next-Generation Optical Communication: Components, Sub-Systems, and Systems III*. Vol. 9009. International Society for Optics and Photonics, Feb. 2014, 90090G.

[110] P. Ebrahimi, M.C. Hauer, Q. Yu, R. Khosravani, D. Gurkan, D.W. Kim, D.W. Lee, and A.E. Willner. "Statistics of Polarization Dependant Gain in Raman Fiber Amplifiers Due to PMD." In: *Technical Digest. Summaries of Papers Presented at the Conference on Lasers and Electro-Optics. Postconference Technical Digest (IEEE Cat. No.01CH37170)*. May 2001, pp. 143–144.

[111] Qiang Lin and Govind P. Agrawal. "Vector Theory of Stimulated Raman Scattering and Its Application to Fiber-Based Raman Amplifiers." In: *Journal of the Optical Society of America B* 20.8 (Aug. 2003), p. 1616.

[112] Andrea Galtarossa, Luca Palmieri, Marco Santagiustina, and Leonora Ursini. "Polarized Backward Raman Amplification in Randomly Birefringent Fibers." In: *Journal of Lightwave Technology* 24.11 (Nov. 2006), pp. 4055–4063.

[113] E. Bettini, A. Galtarossa, L. Palmieri, M. Santagiustina, L. Schenato, and L. Ursini. "Polarized Backward Raman Amplification in Unidirectionally Spun Fibers." In: *IEEE Photonics Technology Letters* 20.1 (Jan. 2008), pp. 27–29.

[114] G. P. Agrawal. *Nonlinear Fiber Optics*. 3rd ed. Optics and Photonics. San Diego: Academic Press, 2001.

[115] Chin-Lin Chen. "Appendix C: Orthogonality and Orthonormality of Guided Modes." In: *Foundations for Guided-Wave Optics*. John Wiley & Sons, Ltd, 2006, pp. 417–420.

[116] Carlo G. Someda. *Electromagnetic Waves*. 2nd ed. Boca Raton, FL: CRC/Taylor & Francis, 2006.

[117] D. Marcuse. "Coupled-Mode Theory for Anisotropic Optical Waveguides." In: *Bell System Technical Journal* 54.6 (July 1975), pp. 985–995.

[118] Robert Hellwarth, Joel Cherlow, and Tien-Tsai Yang. "Origin and Frequency Dependence of Nonlinear Optical Susceptibilities of Glasses." In: *Physical Review B* 11.2 (Jan. 1975), pp. 964–967.

[119] S. Trillo and S. Wabnitz. "Parametric and Raman Amplification in Birefringent Fibers." In: *Journal of the Optical Society of America B* 9.7 (July 1992), p. 1061.

[120] N. Imoto, N. Yoshizawa, J. Sakai, and H. Tsuchiya. "Birefringence in single-mode optical fiber due to elliptical core deformation and stress anisotropy." In: *IEEE Journal of Quantum Electronics* 16.11 (1980), pp. 1267–1271. DOI: 10.1109/JQE.1980.1070382.

[121] D. Chowdhury and D. Wilcox. "Comparison between optical fiber birefringence induced by stress anisotropy and geometric deformation." In: *IEEE Journal of Selected Topics in Quantum Electronics* 6.2 (2000), pp. 227–232. DOI: 10.1109/2944.847757.

[122] P. K. A. Wai and C. R. Menyuk. "Polarization Decorrelation in Optical Fibers with Randomly Varying Birefringence." In: *Optics Letters* 19.19 (Oct. 1994), p. 1517.

[123] Keang-Po Ho and Joseph M. Kahn. "Linear Propagation Effects in Mode-Division Multiplexing Systems." In: *Journal of Lightwave Technology* 32.4 (2014), pp. 614–628. DOI: 10.1109/JLT.2013.2283797.

[124] Filipe Marques Ferreira, Christian S. Costa, Stylianos Sygletos, and Andrew D. Ellis. "Semi-Analytical Modelling of Linear Mode Coupling in Few-Mode Fibers." In: *Journal of Lightwave Technology* 35.18 (Sept. 2017), pp. 4011–4022.

[125] Shaival Buch, Sami Mumtaz, René-Jean Essiambre, Antonia M. Tulino, and Govind P. Agrawal. "Averaged Nonlinear Equations for Multimode Fibers Valid in All Regimes of Random Linear Coupling." In: *Optical Fiber Technology* 48 (Mar. 2019), pp. 123–127.

[126] Brian Hall. *Lie groups, Lie algebras, and representations : an elementary introduction*. Cham: Springer, 2015. ISBN: 978-3-319-13467-3.

[127] Paolo Serena, Chiara Lasagni, Simone Musetti, and Alberto Bononi. "On Numerical Simulations of Ultra-Wideband Long-haul Optical Communication Systems." In: *Journal of Lightwave Technology* (2019), pp. 1–1.

[128] O.V. Sinkin, R. Holzlohner, J. Zweck, and C.R. Menyuk. "Optimization of the Split-Step Fourier Method in Modeling Optical-Fiber Communications Systems." In: *Journal of Lightwave Technology* 21.1 (Jan. 2003), pp. 61–68.

[129] Xueming Liu and Byoungho Lee. "A Fast Method for Nonlinear Schrodinger Equation." In: *IEEE Photonics Technology Letters* 15.11 (Nov. 2003), pp. 1549–1551.

[130] C. R. Menyuk and B. S. Marks. "Interaction of Polarization Mode Dispersion and Nonlinearity in Optical Fiber Transmission Systems." In: *Journal of Lightwave Technology* 24.7 (July 2006), pp. 2806–2826.

[131] COMSOL AB, Stockholm, Sweden. *COMSOL Multiphysics® v. 5.6*. 2019. URL: www.comsol.com.

[132] Luca Palmieri and Andrea Galtarossa. "Intramodal Dispersion Properties of Step-Index Few-Mode Spun Fibers." In: *Journal of Lightwave Technology* 34.2 (Jan. 2016), pp. 303–313.

[133] Liyao Yu, Jian Zhao, Qi Mo, and Guifang Li. "The Beat-Length of Polarization-Maintaining Few-Mode-Fiber Measurement Based on Polarized Interference." In: *2016 15th International Conference on Optical Communications and Networks (ICOCN)*. Sept. 2016, pp. 1–3.

[134] J. M. Dudley and J. R. Taylor, eds. *Supercontinuum Generation in Optical Fibers*. Cambridge ; New York: Cambridge University Press, 2010.

[135] R. R. Alfano and S. L. Shapiro. "Emission in the Region 4000 to 7000 Å Via Four-Photon Coupling in Glass." In: *Physical Review Letters* 24.11 (Mar. 1970), pp. 584–587.

[136] RR Alfano and SL Shapiro. "Picosecond spectroscopy using the inverse Raman effect." In: *Chemical Physics Letters* 8.6 (1971), pp. 631–633.

[137] Joel I. Gersten, R. R. Alfano, and Milivoj Belic. "Combined Stimulated Raman Scattering and Continuum Self-Phase Modulations." In: *Physical Review A* 21.4 (Apr. 1980), pp. 1222–1224.

[138] Akira Hasegawa and Frederick Tappert. "Transmission of Stationary Nonlinear Optical Pulses in Dispersive Dielectric Fibers. I. Anomalous Dispersion." In: *Applied Physics Letters* 23.3 (Aug. 1973), pp. 142–144.

[139] L. F. Mollenauer, R. H. Stolen, and J. P. Gordon. "Experimental Observation of Picosecond Pulse Narrowing and Solitons in Optical Fibers." In: *Physical Review Letters* 45.13 (Sept. 1980), pp. 1095–1098.

[140] F. M. Mitschke and L. F. Mollenauer. "Discovery of the Soliton Self-Frequency Shift." In: *Optics Letters* 11.10 (Oct. 1986), pp. 659–661.

[141] A. Hasegawa and Y. Kodama. "Signal Transmission by Optical Solitons in Monomode Fiber." In: *Proceedings of the IEEE* 69.9 (Sept. 1981), pp. 1145–1150.

[142] H. Northrup, A. L. Beaudet, W. E. O'Brien, G. E. Herman, R. A. Lewis, and M. S. Pollack. "Linkage of Tuberous Sclerosis to ABO Blood Group." In: *Lancet (London, England)* 2.8562 (Oct. 1987), pp. 804–805.

[143] T. A. Birks, J. C. Knight, and P. St J. Russell. "Endlessly Single-Mode Photonic Crystal Fiber." In: *Optics Letters* 22.13 (July 1997), pp. 961–963.

[144] N. G. R. Broderick, T. M. Monro, P. J. Bennett, and D. J. Richardson. "Nonlinearity in Holey Optical Fibers: Measurement and Future Opportunities." In: *Optics Letters* 24.20 (Oct. 1999), pp. 1395–1397.

[145] Jinendra K. Ranka, Robert S. Windeler, and Andrew J. Stentz. "Visible Continuum Generation in Air–Silica Microstructure Optical Fibers with Anomalous Dispersion at 800 Nm." In: *Optics Letters* 25.1 (Jan. 2000), pp. 25–27.

[146] P. Domachuk, N. A. Wolchover, M. Cronin-Golomb, A. Wang, A. K. George, C. M. B. Cordeiro, J. C. Knight, and F. G. Omenetto. "Over 4000 Nm Bandwidth of Mid-IR Supercontinuum Generation in Sub-Centimeter Segments of Highly Nonlinear Tellurite PCFs." In: *Optics Express* 16.10 (May 2008), pp. 7161–7168.

[147] G. Humbert et al. "Supercontinuum Generation System for Optical Coherence Tomography Based on Tapered Photonic Crystal Fibre." In: *Optics Express* 14.4 (Feb. 2006), pp. 1596–1603.

[148] Th Udem, R. Holzwarth, and T. W. Hänsch. "Optical Frequency Metrology." In: *Nature* 416.6877 (Mar. 2002), pp. 233–237.

[149] John M. Dudley, Goëry Genty, and Stéphane Coen. "Supercontinuum Generation in Photonic Crystal Fiber." In: *Reviews of Modern Physics* 78.4 (Oct. 2006), pp. 1135–1184.

[150] Alexander M. Heidt. "Pulse Preserving Flat-Top Supercontinuum Generation in All-Normal Dispersion Photonic Crystal Fibers." In: *Journal of the Optical Society of America B* 27.3 (Mar. 2010), p. 550.

[151] Govind P. Agrawal. "Nonlinear Fiber Optics: Its History and Recent Progress [Invited]." In: *JOSA B* 28.12 (Dec. 2011), A1–A10.

[152] R. H. Stolen, W. J. Tomlinson, H. A. Haus, and J. P. Gordon. "Raman Response Function of Silica-Core Fibers." In: *Journal of the Optical Society of America B* 6.6 (June 1989), p. 1159.

[153] Q. Lin and Govind P. Agrawal. "Raman Response Function for Silica Fibers." In: *Optics Letters* 31.21 (Nov. 2006), p. 3086.

[154] Aleksei Shabat and Vladimir Zakharov. "Exact Theory of Two-Dimensional Self-Focusing and One-Dimensional Self-Modulation of Waves in Nonlinear Media." In: *Soviet physics JETP* 34.1 (1972), p. 62.

[155] James M Stone and Jonathan C Knight. "Visibly "White" Light Generation in Uniform Photonic Crystal Fiber Using a Microchip Laser." In: *Optics express* 16.4 (2008), pp. 2670–2675.

[156] Yuan Liu, Haohua Tu, and Stephen A. Boppart. "Wave-Breaking-Extended Fiber Supercontinuum Generation for High Compression Ratio Transform-Limited Pulse Compression." In: *Optics Letters* 37.12 (June 2012), pp. 2172–2174.

[157] Alexander M. Heidt, Joanna Modupeh Hodasi, Anupamaa Rampur, Dirk-Mathys Spangenberg, Manuel Ryser, Mariusz Klimczak, and Thomas Feurer. "Low Noise All-Fiber Amplification of a Coherent Supercontinuum at 2 *M*m and Its Limits Imposed by Polarization Noise." In: *Scientific Reports* 10.1 (Dec. 2020), p. 16734.

[158] S. Dupont, Z. Qu, S.-S. Kiwanuka, L. E. Hooper, J. C. Knight, S. R. Keiding, and C. F. Kaminski. "Ultra-High Repetition Rate Absorption Spectroscopy with Low Noise Supercontinuum Radiation Generated in an All-Normal Dispersion Fibre." In: *Laser Physics Letters* 11.7 (May 2014), p. 075601.

[159] Alexander Hartung, Alexander M. Heidt, and Hartmut Bartelt. "Design of All-Normal Dispersion Microstructured Optical Fibers for Pulse-Preserving Supercontinuum Generation." In: *Optics Express* 19.8 (Apr. 2011), pp. 7742–7749.

[160] Xin Jiang et al. "Supercontinuum Generation in ZBLAN Glass Photonic Crystal Fiber with Six Nanobore Cores." In: *Optics Letters* 41.18 (Sept. 2016), pp. 4245–4248.

[161] Alexander Hartung, Jörg Bierlich, Adrian Lorenz, Jens Kobelke, and Matthias Jäger. "Design and Fabrication of All-Normal Dispersion Nanohole Suspended-Core Fibers." In: *Journal of the Optical Society of America B* 36.12 (Dec. 2019), p. 3404.

[162] M. A. Eftekhar, L. G. Wright, M. S. Mills, M. Kolesik, R. Amezcua Correa, F. W. Wise, and D. N. Christodoulides. "Versatile Supercontinuum Generation in Parabolic Multimode Optical Fibers." In: *Optics Express* 25.8 (Apr. 2017), pp. 9078–9087.

[163] Francesco Poletti and Peter Horak. "Dynamics of Femtosecond Supercontinuum Generation in Multimode Fibers." In: *Optics Express* 17.8 (Apr. 2009), pp. 6134–6147.

[164] Uğur Teğin and Bülend Ortaç. "Cascaded Raman Scattering Based High Power Octave-Spanning Supercontinuum Generation in Graded-Index Multimode Fibers." In: *Scientific Reports* 8.1 (Aug. 2018), p. 12470.

[165] G. Lopez-Galmiche, Z. Sanjabi Eznaveh, M. A. Eftekhar, J. Antonio Lopez, L. G. Wright, F. Wise, D. Christodoulides, and R. Amezcua Correa. "Visible Supercontinuum Generation in a Graded Index Multimode Fiber Pumped at 1064 Nm." In: *Optics Letters* 41.11 (June 2016), pp. 2553–2556.

[166] Irnis Kubat and Ole Bang. "Multimode Supercontinuum Generation in Chalcogenide Glass Fibres." In: *Optics Express* 24.3 (Feb. 2016), p. 2513.

[167] Ameni Ben Khalifa, Amine Ben Salem, and Rim Cherif. "Mid-Infrared Supercontinuum Generation in Multimode As_2 Se_3 Chalcogenide Photonic Crystal Fiber." In: *Applied Optics* 56.15 (May 2017), pp. 4319–4324.

[168] Alexander Hartung, Jörg Bierlich, Jens Kobelke, and Matthias Jäger. "A New Technique for Efficient Input Coupling into Sub-Wavelength Diameter Suspended Core Fibers." In: *Advanced Photonics 2018 (BGPP, IPR, NP, NOMA, Sensors, Networks, SPPCom, SOF) (2018), Paper SoW3H.6*. Optical Society of America, July 2018, SoW3H.6.

[169] Alexander M. Heidt, James S. Feehan, Jonathan H. V. Price, and Thomas Feurer. "Limits of Coherent Supercontinuum Generation in Normal Dispersion Fibers." In: *Journal of the Optical Society of America B* 34.4 (Apr. 2017), p. 764.

[170] Anupamaa Rampur, Dirk-Mathys Spangenberg, Benoît Sierro, Pascal Hänzi, Mariusz Klimczak, and Alexander M. Heidt. "Perspective on the next Generation of Ultra-Low Noise Fiber Supercontinuum Sources and Their Emerging Applications in Spectroscopy, Imaging, and Ultrafast Photonics." In: *Applied Physics Letters* 118.24 (June 2021), p. 240504.

[171] U. Møller, S. T. Sørensen, C. Jakobsen, J. Johansen, P. M. Moselund, C. L. Thomsen, and O. Bang. "Power Dependence of Supercontinuum Noise in Uniform and Tapered PCFs." In: *Optics Express* 20.3 (Jan. 2012), pp. 2851–2857.

[172] U. Møller and O. Bang. "Intensity Noise in Normal-Pumped Picosecond Supercontinuum Generation, Where Higher-Order Raman Lines Cross into Anomalous Dispersion Regime." In: *Electronics Letters* 49.1 (2013), pp. 63–65.

[173] Anupamaa Rampur et al. "Ultra Low-Noise Coherent Supercontinuum Amplification and Compression below 100 Fs in an All-Fiber Polarization-Maintaining Thulium Fiber Amplifier." In: *Optics Express* 27.24 (Nov. 2019), pp. 35041–35051.

[174] Ronen Dar, Meir Feder, Antonio Mecozzi, and Mark Shtaif. "Properties of Nonlinear Noise in Long, Dispersion-Uncompensated Fiber Links." In: *Optics Express* 21.22 (Nov. 2013), p. 25685.

[175] Cristian Antonelli, Ori Golani, Mark Shtaif, and Antonio Mecozzi. "Nonlinear Interference Noise in Space-Division Multiplexed Transmission through Optical Fibers." In: *Optics Express* 25.12 (June 2017), pp. 13055–13078.

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede and Ivo Pletikosić. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/

Thank you very much for your feedback and contribution.