Stella Civelli

Nonlinear frequency-division multiplexing: theoretical aspects, numerical algorithms, and experimental demonstration

Sant'Anna
School of Advanced Studies – Pisa

Academic Year 2017/2018
Phd Course in Emerging Digital Technologies

Nonlinear frequency-division multiplexing: theoretical aspects, numerical algorithms, and experimental demonstration

Author
Stella Civelli

Supervisor
Marco Secondini

Se puoi vedere, guarda.

Se puoi guardare, osserva.

Libro dei Consigli.

Gli errori sono necessari,

utili come il pane e spesso anche belli:

per esempio la torre di Pisa.

Gianni Rodari.

Acknowledgments

I am profoundly and sincerely grateful to Dr. Marco Secondini for supervising my work with curiosity and attention. He shared honest advice and brilliant ideas with me; he gave me freedom in my research path, but always supervised my work with accuracy. Marco always encouraged me to new challenges, showing me trust and confidence in my abilities.

I would like to thank Prof. Enrico Forestieri for his valuable help and useful suggestions. Also, I am very thankful to all my colleagues from the TeCIP Institute, who have always been funny and helpful, and made my journey enjoyable and smooth.

I am very grateful to Prof. Sergei K. Turitsyn and Dr. Jaroslaw E. Prilepsky for hosting me at the Aston Institute of Photonic Technologies, Aston University, in Birmingham (UK) for six months. I believe that our collaboration produced good results, and I hope that it will continue in the future. I am thankful to all the institute, that made my stay very nice.

Also, I would like to thank Prof. Darko Zibar and all the Photonic Department of the Technical University of Denmark (DTU) that hosted me for a very short but intense period. I sincerely enjoyed my stay at DTU. In particular, I would like to thank Dr. Francesco Da Ros and Dr. Simone Gaiarin for their friendly attitude during our hard work.

I thank the Electromagnetics Academy and The Rank Prize Funds for sponsoring two prizes that I have won for my research: the Best Student Paper Award (2nd prize) at the PIERS Conference 2018 and the Best Contributed Paper at the Symposium on Challenges to Achieving Capacity in Nonlinear Optical Networks 2018, respectively.

Finally, I am extremely grateful to my family, my better half, and my friends. This would not have been possible without their support, their teachings, and their guidance.

Abstract

The exponential growth of global data traffic is challenging the optical network, currently limited by Kerr nonlinearity. This thesis investigates a novel transmission paradigm for long haul optical fiber communication, expected to outperform conventional systems. This technique, commonly known as nonlinear frequency-division multiplexing (NFDM), uses the nonlinear Fourier transform (NFT) to encode information on the nonlinear spectrum, whose evolution along the optical fiber is, under some circumstances, trivial. For this reason, NFDM masters nonlinearity, and is expected to increase the capacity of the current optical networks. The first part of this thesis is dedicated to the NFT as a tool for optical communication: the mathematical framework is introduced and numerical algorithms for the NFT are developed. Secondly, the implementation of an NFDM scheme is described and the performance, obtained through simulations, is shown. The results are discussed and some important limitations are highlighted. Next, in the light of the above, three novel detection strategies are introduced that provide significant performance improvements (up to 6.2 dB). Then, the main concepts are extended to dual-polarization systems, and a reduced complexity paradigm is introduced. Finally, the first experimental demonstration of a dual-polarization NFDM scheme modulating both the continuous and the discrete spectra is described and the results are shown.

Contents

List of acronyms

ADC analog-to-digital converter

AWG arbitrary waveform generator

AWGN additive white gaussian noise

AOM acusto-optic modulator

ASE amplified spontaneous emission

B2B back-to-back

BER bit error rate

BNFT backward NFT

BPD balanced photodetector

BPS blind phase search

DAC digital-to-analog converter

DBP digital backpropagation

DF-BNFT decision-feedback BNFT

DF-FNFT decision-feedback FNFT

DSO digital storage oscilloscope

DSP digital signal processing

DT Darboux transform

EDC electronic dispersion compensation

EDFA erbium-doped fiber amplifier

EVM error vector magnitude

FFT fast Fourier transform

FNFT forward NFT

FT Fourier transform

FWHM full width at half maximum

FWM four-wave mixing

GLME Gelfand-Levitan-Marchenko equation

GVD group velocity dispersion

HD-FEC hard decision forward error correction

KdV Korteweg-de Vries

I-FNFT incremental FNFT

IQ in-phase and quadrature

ISI intersymbol interference

ISO isolator

IST inverse scattering transform

IVP initial value problem

LODE linear ordinary differential equation

LP Layer-Peeling

LPA lossless path-averaged

MAP maximum a posteriori probability

ME Manakov equation

NADM nonlinear add-drop multiplexer

NCG Nyström-conjugate gradient

NIS nonlinear inverse synthesis

NFDM nonlinear frequency-division multiplexing

NFT nonlinear Fourier transform

NLSE nonlinear Schrödinger equation

NMSE normalized mean square error

OBPF optical band pass filter

OFDM orthogonal frequency-division multiplexing

OPC optical phase conjugation

OSNR optical signal-to-noise ratio

PAPR peak-to-average power ratio

PC polarization controller

pdf probability density function

PDM-NIS polarization-division multiplexing NIS

PMD polarization mode dispersion

PNFDM polarization and nonlinear frequency-division multiplexing

LO local oscillator

PRBS pseudo-random bit sequence

PSD power spectral density

QAM quadrature amplitude modulation

QPSK quadrature phase-shift keying

RC raised cosine

ROADM reconfigurable optical add-drop multiplexer

RRC root raised cosine

RX receiver

SDM space division multiplexing

SE spectral efficiency

SMF single mode fiber

SNR signal-to-noise ratio

SSFM split step Fourier method

SPM self-phase modulation

TX transmitter

VGLME vector GLME

VNLSE vector NLSE

WDM wavelength-division multiplexing

XPM cross-phase modulation

Z-S Zakharov-Shabat

List of Figures

Chapter 1

Introduction

This thesis investigates a pioneering transmission paradigm for long haul optical fiber communication, which aims to master the inherent nonlinearity of the optical fiber channel. Starting from a mathematical concept—the nonlinear Fourier transform (NFT) or inverse scattering transform (IST)—the theoretical framework and corresponding numerical algorithms are developed; the corresponding communication system—the nonlinear frequency-division multiplexing (NFDM)—is designed and analyzed. The performance of different flavors of NFDM are investigated through simulations and discussed. Finally, an NFDM transmission using all degrees of freedom (both polarizations and spectrum) is experimentally demonstrated.

This Chapter is organized as follows. Section 1.1 explains the motivation behind this work. The contributions of this thesis, as well as my own contribution to the overall work are detailed in Section 1.2. Finally, Section 1.3 describes the thesis structure.

1.1 Motivations

The history of telecommunication goes back to the use of smoke signals and drums. Since the first half of the XIX century, when the telegraph was invented, the advancing in telecommunication has been unstoppable. The exponential growth of the bit rate-distance product, a typical performance metric in telecommunication, was enabled by the development of novel technologies, as shown in Fig. 1.1 [1]. Particularly, the invention of the laser in the 1960s permitted the advent of lightwave communication, which, thanks to the enormous available bandwidth, boosted the performance of communication systems [1]. In 1970, after the work of C. Kao and G. Hockam [2], the first optical fiber with sufficient low loss was developed [1, 3]. Fiber optic communication systems, i.e., lightwave systems that use optical fibers as a mean to transmit information, have been deployed since 1980 [1, 4]. Then, the advent of optical amplifiers allowed to mitigate the limits imposed by loss in optical fibers, enabling longer

distances without regenerating the signal. Next, wavelength-division multiplexing (WDM) and higher order modulation formats allowed to further increase the bit rate-distance product.

Nowadays, the global data traffic is increasing exponentially, constantly challenging the capability of current-generation optical fiber communication systems to meet the data rate demand [3–5]. The nonlinearity of optical fiber systems is believed to be the main limiting factor deteriorating the performance at high signal powers [4, 5], and, thus, currently limiting the throughput of single mode fiber (SMF). To address the future capacity needs of optical fiber networks and forestall the infamous "capacity crunch" problem [4], two solutions have been widely considered: space division multiplexing (SDM) [3], to increase the throughput of a communication system by multiplexing information in space, and various nonlinearity compensation techniques to enhance the throughput of SMF, i.e., on each single spatial dimension. On the one hand, SDM would require either the installation of new multi-mode or multi-core fibers in place of current-generation fibers, or simply to increase the number of conventional SMF. Both approaches have to face serious problems in terms of deployment costs. On the other hand, because of the huge number of already installed fibers and the obvious engineers' goal to maximize the information rate for every available spatial dimension (fiber, core, or mode), there exists a great interest in the area of compensation, mitigation [6], or constructive use of fiber nonlinearity [7]. More specifically, optical fiber communication systems have been developed extending concept from linear communication (e.g., wireless or radio) and essentially treated the optical fiber as a linear channel. However, the demand for higher bit rate-distance product progressively increased the launch power in fibers, bringing in nonlinear effects. Consequently, a lot of effort has been done to compensate and mitigate nonlinear effects. A brief discussion about nonlinearity compensation and mitigation techniques is reported in Section 2.3. Conversely, in the past years, a revolutionary approach has been actively investigated to use fiber nonlinearity in a constructive way, rather than treating it as a perturbation. This approach uses the NFT [8–10]—a sort of nonlinear analog of the conventional Fourier transform (FT)—to decompose a signal into a set of discrete and continuous spectral components, the so-called nonlinear spectrum, that evolves in a simple linear way along the nonlinear fiber channel. NFDM [7, 11–19, and reference therein] is the umbrella name for optical fiber transmission techniques which encode the information on the nonlinear spectrum, such that, differently from conventional WDM, the different users are assigned to different domains ("bands") in the NFT spectrum. The latter evolves linearly along the fiber, which guarantees the absence of crosstalk between users (responsible for a severe performance degradation in WDM systems) and to exactly remove propagation effects in a single tap operation[1]. These characteristics make the NFDM

[1]Despite the presence of solitons, NFDM is not soliton communication [20, 21], and is, instead, rather different. Just to mention a few important differences: NFDM can use quadrature amplitude modulation (QAM) constellations, include the modulation of the continuous nonlinear spectrum, and master the interactions between "modes" i.e., continuous-discrete spectra interactions and discrete-discrete spectra interactions.

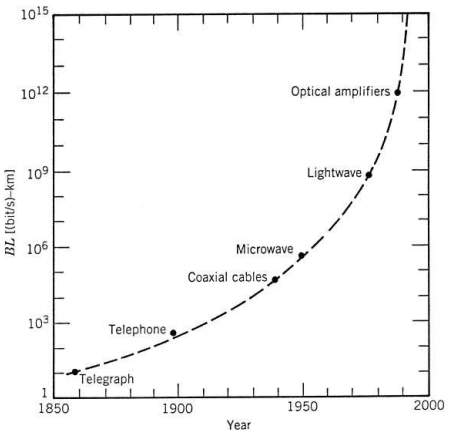


Figure 1.1: The history of telecommunications. Image taken from [1].

a good candidate for the next generation of fiber systems, taking into account inherent robustness to fiber nonlinearity and the potential to outperform conventional "nonlinearity-degraded" systems [7,11,12,18,22]. Moreover, the recent observation of solitons in conventional communication formats—i.e., orthogonal frequency-division multiplexing (OFDM) signals—further encourages the use of the NFT in optical communication [23]. Finally, it is worth noting that NFDM does not exclude SDM, instead, the use of SDM-NFDM is feasible and deserves investigation (see Section 5.3).

To conclude, I would like to cite the notice on the 2006 Steele Prizes regarding the Seminal Contribution to Research prize awarded to Clifford S. Gardner, John M. Greene, Martin D. Kruskal, and Robert M. Miura for their fundamental paper in the NFT theory "nonlinearity has undergone a revolution: from a nuisance to be eliminated to a new tool to be exploited" [24, 25].

1.2 Contributions

The subject of this thesis is NFDM—a pioneering transmission technique for long haul optical fiber communication which aims at mastering nonlinearity and achieving better performance with respect to conventional systems. NFDM is a very specific as well as interdisciplinary topic, which requires knowledge in information theory, optical fiber communications, and mathematics. Despite the fact that many groups recently started working on NFDM, there are still many unknowns and open questions in the field, due to the pioneering nature of the topic as well as the complexity and nonlinear nature of the mathematical theory behind. The aim of this thesis is to give a contribution through theoretical and simulation-based analysis for the understanding and optimization of the scheme, with a medium-long term perspective. For more details about the contributions

and outcomes of this thesis, the reader can refer to the final Chapter 7.

As far as it concerns the development of the work, I have had a major contribution in all the work presented in this thesis. As far as it concerns Chapters 3, 4, and 5, I have developed some of the theoretical properties related to the NFT and the numerical algorithms, I performed the simulations, and I analyzed and discussed the results under the supervision of M. Secondini and with the collaboration of E. Forestieri for single-polarization related topics, and Y. E. Prilepsky and S. K. Turitsyn for dual-polarization related topics. As far as it concerns Chapters 6, my contribution primarily concerned (together with all the authors of [26] and, in particular, F. Da Ros and S. Gaiarin) the development of the numerical joint backward NFT (BNFT), the design of the scheme, and the discussion of the results; concerning the experimental part, I participated in the experiment (performed at the Technical University of Denmark), tough I did not directly performed it.

1.3 Thesis outline

This thesis is divided into three main parts, which reflect both the logical and the chronological order in which the research have been developed: mathematical and numerical framework, NFDM analysis through simulations, and NFDM experimental demonstration.

Chapter 2 contains some useful preliminary information. Chapter 3 deals with the NFT for the scalar and vector nonlinear Schrödinger equation (NLSE), with particular attention to the application to optical fiber communication. The mathematical framework is presented and some properties are elucidated; next, numerical algorithms for the forward (direct) and backward (inverse) NFT are described. Chapter 4 investigates single-polarization NFDM through theoretical analysis and simulations, with particular attention to different modulation schemes and improved detection strategies. Chapter 5 extends some of the concepts of Chapter 4 to dual-polarization schemes. Chapter 6 concerns the first experimental demonstration of dual-polarization NFDM transmission with joint (continuous and discrete) spectrum modulation. Finally, Chapter 7 draws the conclusions and possible outlooks for future works.

Chapter 2

Preliminaries

This chapter very briefly resumes some well established concepts in information theory and optical fiber communication in Sections 2.1 and 2.2, respectively, and provides an overview of the most common nonlinearity compensation techniques in Section 2.3. Finally, Section 2.4 contains some notations and useful definitions that are used thorough this thesis.

2.1 Capacity of a channel

The bit rate R_b is the number of bits that are transmitted in one second and is measured in bits per second. The capacity of a (memoryless discrete-time) channel is defined as the maximum over the input distributions of the mutual information between the input and the output. Remarkably, the maximum bit rate that can be transmitted through a channel with an arbitrary low error probability is the capacity C of the channel [27]. Furthermore, the capacity of an additive white gaussian noise (AWGN) channel is [27]

$$C = B \log_2(1 + \mathrm{SNR}), \text{ in bit/s} \tag{2.1}$$

where B is the channel bandwidth in Hz and the signal-to-noise ratio (SNR) is the ratio between the average signal power P and the average noise power. If the noise power spectral density (PSD) is $N_0/2$, then $\mathrm{SNR} = P/(N_0 B)$ [28].

2.2 Optical fiber communication

A communication system is designed to transfer information from one point to another. Optical communication systems—or lightwave systems—are communication systems that uses an optical signal (visible or near-infrared region of the electromagnetic spectrum $\sim 100\,\mathrm{THz}$) to carry information. Such an high carrier frequency allows a potential enormous bandwidth. An optical fiber communication system is an optical communication system that uses an optical

fiber as transmission channel [1]. A typical WDM optical fiber communication system is sketched in Fig. 2.1. Each transmitter (TX) encodes the information on a signal with a specific wavelength; the signals are multiplexed together and sent into the channel. The channel is made of several spans of optical fibers and optical amplifiers. At the output of the channel, the signals are demultiplexed and the information is retrieved by a bank of receivers (RXs).

Using optical fibers as a medium has many advantages (very low weight and size, flexibility, immunity to electromagnetic interference, etc.). There are, however, several effects that need to be taken into account when designing an optical fiber transmission scheme. First of all, optical fibers are lossy, i.e., the power of the optical signal decreases during propagation (with a typical attenuation coefficient α of 0.2 dB/km) [1]. Therefore, to achieve longer distances it is essential to amplify the optical signal using optical amplifiers[1]. Unfortunately, optical amplifiers introduce ASE noise, which can be modeled as AWGN. The overall effect of an amplifier on a signal $s(t)$ is depicted in Fig. 2.2, where G is the amplifier gain and $n(t)$ is AWGN. Another possible solution is distributed Raman amplification, which amplifies the signal along the fiber link. Distributed Raman amplification also causes ASE noise. Once loss is accounted for, dispersion becomes the most relevant effect in optical fibers. Dispersion is caused by the fact that different frequencies travel at different speed during propagation, therefore causing temporal broadening of the signal [1]. Different dispersion compensation techniques can be considered, including dispersion-compensating fibers, low-dispersion fibers, and electronic dispersion compensation (EDC). Dispersion is usually described by the group velocity dispersion (GVD) parameter β_2, which is measured in ps^2/km, or by the dispersion parameter $D = -2\pi c\beta_2/\lambda^2$ in $\text{ps}/(\text{nm}\cdot\text{km})$, with c being the speed of light and λ the considered wavelength.

Once both loss and dispersion effects are compensated for, Kerr nonlinearity becomes the limiting factor. Indeed, according to Eq. (2.1), one may try to increase the signal power (and, therefore, the SNR) to reduce the impact of noise and achieve a higher bit rate, but Kerr nonlinearity becomes relevant. The Kerr effect [29] is the change in refractive index n caused by the intensity I of the optical signal as $n = n_0 + n_2 I$, where n_2 is the nonlinear refractive index. All these three effects are described by the well known NLSE (2.3), described in the following. Another important effect that occurs in SMF is PMD: due to inhomogeneities, the conventional SMFs are birefringent and support two orthogonal modes that can generally have a different group velocity [29]. We will return on this later in this Section.

2.2.1 The nonlinear Schrödinger equation

Light propagation in optical fiber is described by the wave equation

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E} = -\frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} - \frac{1}{\mu_0}\frac{\partial^2 \mathbf{P}}{\partial t^2} \tag{2.2}$$

[1] One might also regenerate the signal, but this would require to receive and re-transmit the signal for each wavelength, while amplifiers amplify the multiplexed signal in the optical domain, i.e., without any conversion to electric signals

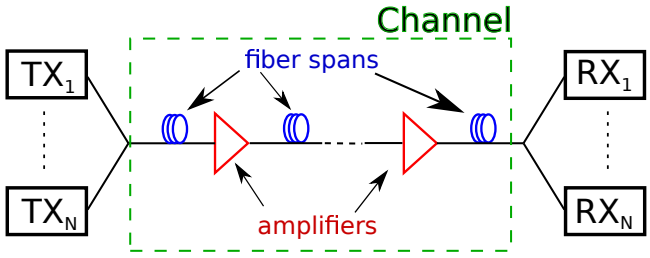


Figure 2.1: A typical optical fiber communication system.

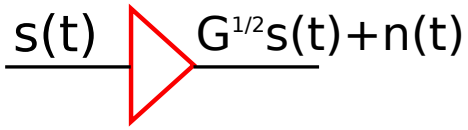


Figure 2.2: An optical amplifier.

for the electric field $\mathbf{E}$, and induced electric polarization $\mathbf{P}$, where c is the speed of light in vacuum and μ_0 is the vacuum permeability [29].

Under the assumptions that (i) the nonlinear part $\mathbf{P}_{NL}$ of $\mathbf{P}$ is small compared to the linear counterpart $\mathbf{P}_L$, (ii) polarization is maintained along the fiber, (iii) the pulse is quasi-monochromatic, and (iv) higher-order dispersion and nonlinear effects can be neglected, one can derive the equation describing the propagation of the complex envelope (See Appendix A.1) $q = q(z,t)$ of the electric field $\mathbf{E}$ [7, 29]

$$j\frac{\partial q}{\partial z} - \frac{\beta_2}{2}\frac{\partial^2 q}{\partial t^2} + \gamma|q|^2 q + j\frac{\alpha}{2}q = n, \tag{2.3}$$

where z is the space coordinate along the fiber, t is the retarded time moving with the group velocity, and $n = n(z,t)$ the amplified spontaneous emission (ASE) noise caused by amplification. The second term accounts for dispersion and causes pulse broadening during propagation, while the third term accounts for Kerr nonlinearity, which is responsible for several effects including self-phase modulation (SPM), cross-phase modulation (XPM), and four-wave mixing (FWM). Finally, the fourth term accounts for loss during propagation.

The term responsible for loss can be removed considering ideal distributed Raman amplification. Otherwise, it can be neglected with good approximation considering the lossless path-averaged (LPA) [30] model, i.e., replacing the nonlinear coefficient γ with its average value over all the fiber (of length L)

$$\gamma_1 = \gamma\frac{1}{L}\int_0^L A(z)^{-1}\,dz, \tag{2.4}$$

where $A(z)$ is the attenuation $A(z) = P(z)/P(0)$, $P(z)$ being the optical power. In particular, when there are equally spaced amplifiers and the loss is exactly

compensated for by the gain $G = e^{\alpha L}$, the average value in (2.4) becomes

$$\gamma_1 = \gamma(G-1)/(G\ln(G)). \tag{2.5}$$

In the noise-free and lossless case, Eq. (2.3) reduces to the NLSE

$$j\frac{\partial q}{\partial z} = \frac{\beta_2}{2}\frac{\partial^2 q}{\partial t^2} - \gamma|q|^2 q. \tag{2.6}$$

2.2.2 The Manakov equation

The assumption about constant polarization considered in the derivation of the previous section is not realistic for conventional SMF. Indeed, it is well known that due to inhomogeneities, the conventional SMFs are birefringent and support two orthogonal modes that can generally have a different group velocity. Birefringence randomly varies both in magnitude and direction along the fiber, causing a phenomenon known as polarization mode dispersion (PMD) [29]. Averaging over the rapidly varying birefringence yields the Manakov-PMD equation [31, 32] for two-component electric field envelope $\mathbf{q} = (q_1, q_2)$

$$j\frac{\partial \mathbf{q}}{\partial z} - \frac{\beta_2}{2}\frac{\partial^2 \mathbf{q}}{\partial t^2} + \gamma\frac{8}{9}||\mathbf{q}||^2\mathbf{q} + j\frac{\alpha}{2}\mathbf{q} = -j\Delta\beta'(z)\bar{\sigma}_3\frac{\partial \mathbf{q}}{\partial t} + \frac{\gamma}{3}\left\{\left[\mathbf{q}^\dagger\bar{\sigma}_2\mathbf{q}\right]\bar{\sigma}_2\mathbf{q} - \frac{1}{3}|\mathbf{q}|^2\mathbf{q}\right\}. \tag{2.7}$$

Without going into the details (the interested reader can refer to [31, 32]), the Manakov-PMD equation separates the slowly varying terms on the left, from the rapidly varying ones on the right. In particular, the first term on the right-hand side of (2.7) represents linear PMD, while the second term represents nonlinear PMD, which is typically negligible in communication systems. Consequently, considering typical optical fibers used in communication systems and neglecting linear PMD effects and loss (i.e., assuming LPA or ideal distributed Raman amplification), the Manakov-PMD equation in the noise-free scenario reduces to the integrable Manakov equation (ME) [31, 33]

$$j\frac{\partial \mathbf{q}}{\partial z} = \frac{\beta_2}{2}\frac{\partial^2 \mathbf{q}}{\partial t^2} - \gamma\frac{8}{9}||\mathbf{q}||^2\mathbf{q}, \tag{2.8}$$

where, again, z is the coordinate along the fiber, and t is the retarded time. Importantly, Eq. (2.8) is a system of two joint scalar NLSE-like equations; the two equations are not independent and their interaction lays in the nonlinear term

$$-\gamma\frac{8}{9}\left(|q_1|^2 + |q_2|^2\right)q_k \text{ for } k = 1, 2. \tag{2.9}$$

Furthermore, note that if the initial pulse is set on a single-polarization as $\mathbf{q} = (q_1, 0)$, the ME reduces further to the NLSE form in (2.6).

2.2.3 The vector nonlinear Schrödinger equation

Interestingly, the ME can be extended to describe the propagation in multi-mode or multi-core fiber in the strong coupling regime [34, 35] yielding the M dimensional vector NLSE (VNLSE)

$$j\frac{\partial \mathbf{q}}{\partial z} = \frac{\beta_2}{2}\frac{\partial^2 \mathbf{q}}{\partial t^2} - \gamma\kappa||\mathbf{q}||^2\mathbf{q}, \tag{2.10}$$

for the M dimensional $\mathbf{q} = \mathbf{q}(z,t)$, where κ accounts for the nonlinear coupling between spatial modes ($\kappa = 8/9$ when $M = 2$ and $\kappa \to 0$ for increasing M). An interested reader can refer to [34, 35] for further information.

2.3 Nonlinearity compensation techniques

Kerr nonlinearity currently limits the performance of optical fiber communication systems. Consequently, numerous techniques, both optical and digital, have been proposed to mitigate and/or compensate for nonlinear effects. Here, we briefly review some of those.

As far as it concerns all-optical compensation techniques, optical phase conjugation (OPC) reverse nonlinear effects by (optically) conjugating the signal in the middle of the link, such that nonlinear effects produced in the second half of the link compensate for those produced in the first half [36]. However, this technique is limited in flexibility by the requirement on the position of the OPC.

Regarding digital signal processing techniques for nonlinearity, numerous methods have been proposed [6], which includes digital backpropagation (DBP), Volterra series-based compensation, and geometric and probabilistic shaping. DBP—the most commonly used technique—consists in undoing deterministic channel effects *digitally* propagating the received signal in an ideal fiber with reversed dispersive and nonlinear effects, i.e., applying the split step Fourier method (SSFM) with opposite sign on β_2 and γ to the samples of the received signal. In this way, one can ideally remove all deterministic signal-signal interactions. However, DBP has some drawbacks. First, in a single-user scenario, many steps should be considered to properly account for the interaction between dispersion and nonlinearity, which significantly increases the computational costs. Moreover, in a multi-user scenario, ideal DBP would require the knowledge of all the other channels, which is both impossible and impracticable from a numerical point of view. Consequently, intra-channel interactions are usually treated as additional noise and, therefore, limit the performance of DBP.

Recently, a revolutionary approach based on the nonlinear spectrum modulation using the NFT has been taken into consideration as a way to master the nonlinearity of the optical fiber channel, rather than using perturbation and approximation-based methods to mitigate it [7, 11, 18, 19]. This approach is the subject of this thesis.

2.4 Notations and useful definitions

Imaginary unit The imaginary unit is denoted as $j = \sqrt{-1}$.

Real and complex number Real and complex number are indicated, respectively, with $\mathbb{R}$ and $\mathbb{C}$. If z is a complex number, its real and imaginary parts are indicated as $\Re(z)$ and $\Im(z)$, respectively. The set $\mathbb{C}^+$ indicates the set of complex numbers with positive imaginary part, i.e., $\mathbb{C}^+ = \{z \in \mathbb{C} : \Im(z) > 0\}$.

Vectors Vectors are indicated with lower case bold characters, and their components with non-bold and subscripts. The vector $\mathbf{v} \in \mathbb{C}^{1\times N}$ is the row vector $\mathbf{v} = (v_1, \ldots, v_N)$ of length N with entries in $\mathbb{C}$; the k-th component of $\mathbf{v}$ is v_k.

Matrices Matrices are indicated with upper case characters, and their components with subscripts. The matrix $A \in \mathbb{C}^{N\times M}$ has dimension $N \times M$ and complex entries; the component with row k and column ℓ is indicated with $A_{k,\ell}$. Moreover, $\mathbf{0}_{N\times M}$ indicates the matrix with all zero entries, while $\mathbf{I}_N$ indicates the identity matrix of dimension $N \times N$. Diagonal matrices having vector $\mathbf{d}$ as diagonal are indicated as $D = D(\mathbf{d})$. Empty spaces in matrices indicate zero entries.

Conjugate The conjugate of a is indicated as a^*, where a can be a scalar, a vector, or a matrix.

Transpose The transpose of the matrix A (or the vector $\mathbf{v}$) is indicated as A^{T} (or $\mathbf{v}^{\mathrm{T}}$).

Conjugate transpose The conjugate and transpose of the matrix A (or the vector $\mathbf{v}$) is indicated as $A^\dagger$ (or $\mathbf{v}^\dagger$) and is equal to $A^\dagger = A^{*\mathrm{T}}$.

Signal A signal is a function $s(t) : \mathbb{R} \to \mathbb{C}$, where t represents time.

Fourier transform (FT) The FT—referred to as spectrum—of the signal $s(t)$ is the function $S(f) : \mathbb{R} \to \mathbb{C}$ defined over the frequencies f as

$$\mathcal{F}\{s(t)\}(f) = S(f) = \int_{-\infty}^{+\infty} s(t)e^{-2\pi jft}dt. \tag{2.11}$$

L^1 space and norm The L^1-norm is defined as

$$||s(t)||_1 = \int_{-\infty}^{+\infty} |s(t)|\, dt. \tag{2.12}$$

If $||s(t)||_1$ is finite, we may say that the signal $s(t)$ is in $L^1(\mathbb{R})$, i.e., $s(t) \in L^1(\mathbb{R})$.

Hankel matrices An upper left triangular Hankel matrix $\mathcal{H}$ of dimension

$N_{\mathcal{H}} \times N_{\mathcal{H}}$, generated by the vector $\mathbf{h} = (h_1, \ldots, h_{N_{\mathcal{H}}})$, is the matrix

$$\mathcal{H} = \mathcal{H}(\mathbf{h}) = \begin{pmatrix} h_1 & h_2 & h_3 & \ldots & h_{N_{\mathcal{H}}} \\ h_2 & h_3 & \ldots & h_{N_{\mathcal{H}}} & 0 \\ h_3 & \ldots & h_{N_{\mathcal{H}}} & & \vdots \\ \ldots & h_{N_{\mathcal{H}}} & & & \\ h_{N_{\mathcal{H}}} & 0 & \ldots & & 0 \end{pmatrix}, \tag{2.13}$$

having $\mathbf{h}$ as first row and $\mathbf{h}^{\mathrm{T}}$ as first column.

Circulant matrices The circulant matrix of dimension $N_{\mathcal{C}} \times N_{\mathcal{C}}$ generated by the vector $\mathbf{c} = (c_1, \ldots, c_{N_{\mathcal{C}}})$ is the matrix

$$\mathcal{C}(\mathbf{c}) = \begin{pmatrix} c_1 & c_2 & c_3 & \ldots & c_{N_{\mathcal{C}}} \\ c_{N_{\mathcal{C}}} & c_1 & c_2 & \ldots & \vdots \\ c_{N_{\mathcal{C}}-1} & c_{N_{\mathcal{C}}} & c_1 & \ldots & c_3 \\ \vdots & \ddots & \ddots & \ddots & c_2 \\ c_2 & \ldots & c_{N_{\mathcal{C}}-1} & c_{N_{\mathcal{C}}} & c_1 \end{pmatrix}, \tag{2.14}$$

having $\mathbf{c}$ as first row and $\tilde{\mathbf{c}} = (c_1, c_{N_{\mathcal{C}}}, \ldots, c_2)^{\mathrm{T}}$ as first column.

$\mathcal{Q}$-function The $\mathcal{Q}$-function is defined as

$$\mathcal{Q}(x) = \int_x^{\infty} e^{-t^2/2}\, dt/\sqrt{2\pi} \tag{2.15}$$

and it is related to the complementary error function $\mathrm{erfc}(\cdot)$ with

$$\mathcal{Q}(x) = \mathrm{erfc}(x/\sqrt{2})/2. \tag{2.16}$$

Rate efficiency The rate efficiency η is defined as the number of information symbols over the total number of symbols (information and guard). Therefore,

$$\eta = \begin{cases} \dfrac{N_b}{N_b + N_z} & \text{single pol.} \\ \\ \dfrac{2N_b}{N_b + N_z} & \text{dual pol.} \end{cases}, \tag{2.17}$$

with N_b being the burst length (i.e., the number of symbols per polarization) and N_z being the guard length (i.e., the number of guard symbols per polarization).

Energy per symbol The energy per information symbol E_s of the signal $\mathbf{q}(t)$ is defined as the total energy over the number of information symbols, and therefore

$$E_s = \begin{cases} (N_b)^{-1} \displaystyle\int_{-\infty}^{+\infty} |q(t)|^2 dt & \text{single pol.} \\ \\ (2N_b)^{-1} \int_{-\infty}^{+\infty} \left(|q_1(t)|^2 + |q_2(t)|^2\right) dt & \text{dual pol.} \end{cases}, \tag{2.18}$$

Power per symbol The average power per information symbol P_s (and polarization) is defined as

$$P_s = E_s/T_s, \tag{2.19}$$

T_s being the symbol time, and E_s the energy per symbol (2.18). The actual average optical power is ηP_s. In this manuscript, except for Chapter 6, the average power per information symbol is considered for performance evaluation.

Q-factor The Q-factor is used to measure performance. The Q^2 is defined as

$$Q^2_{\mathrm{dB}} = 20 \log_{10}[\sqrt{2}\,\mathrm{erfc}^{-1}(2P_b)], \tag{2.20}$$

where the probability of bit error P_b is given by direct error counting [37], or estimated through the error vector magnitude (EVM) [38, 39]. For an M-QAM constellation

$$P_b \sim \frac{2(1-1/\sqrt{M})}{\log_2 M}\,\mathrm{erfc}\left(\frac{1}{\mathrm{EVM}}\sqrt{\frac{3}{2(M-1)}}\right). \tag{2.21}$$

In this manuscript, the Q-factor is evaluated from direct bit error rate (BER) counting in Chapter 6 (since Eq. (2.21) may not be accurate for the discrete spectrum) and for performance evaluation in Section 4.5 (since the Q-factor can not be evaluated in some cases). In all the other cases, Eq. (2.21) is used.

Oversampling factors The oversampling factors for the forward NFT (FNFT) and the BNFT, are indicated with N_F and N_B, respectively.

Chapter 3

The nonlinear Fourier transform

The NFT—most commonly know as IST outside the engineering community—is a mathematical method to solve the initial value problem (IVP) associated with a certain class of nonlinear partial differential equations, referred to as integrable equations. In a nutshell, the method consists in finding the solution of an associated linear ordinary differential equation (LODE) (scattering problem), which describes the solution of the IVP, and whose evolution is given by another linear equation. In simpler words, an integrable nonlinear differential equation can be *s*olved moving to another domain—the nonlinear frequency domain—where it is transformed into a *simple* equation. Consequently, the NFT can be regarded as the nonlinear analogous of the well-know linear FT, which, indeed, approximates the NFT at low powers.

This Chapter is organized as follows. After a brief introduction about the NFT in Section 3.1, Section 3.2 reviews the NFT theory regarding the VNLSE with vanishing boundary conditions, and Section 3.3 highlights some of the concepts regarding the scalar NLSE. Next, Section 3.4 describes some methods for the numerical computation of the NFTs operations for the general multi dimensional case. Finally, Section 3.5 mentions some of the possible applications of the NFT to optical fiber communication, and Section 3.6 draws the conclusions. This Chapter reports some of the numerical methods presented in [26, 40, 41].

3.1 Preliminaries

3.1.1 Brief history of the nonlinear Fourier transform

The history of the NFT goes back to 1834, when J. Scott Russel [42] observed for the first time a solitary wave—a soliton—in water. In 1838 he wrote

"a large solitary elevation, a rounded, smooth and well defined heap of water, which continued its course along the channel apparently without change of form

or diminuition of speed.... Its height gradually diminished, and after a chase of one or two miles I lost it in the windings of the channel. Such, in the month of August 1834, was my first chance interview with that singular and beautiful phenomenon."

In 1895 Korteweg and de Vries derived the Korteweg-de Vries (KdV) equation[1] [11, 43]

$$\frac{\partial q}{\partial z} = q\frac{\partial q}{\partial t} + \frac{\partial^3 q}{\partial t^3} \tag{3.1}$$

for moderately small, shallow-water waves, which admits solitary wave solutions. Later, in 1965 Zabusky and Kruskal [44] showed numerically that these special solutions remain unaffected, except for a phase shift, after interaction with each others, and named these solutions *solitons*.

In 1967, Gardner, Greene, Kruskal and Miura [45] presented a method to solve the KdV equation with an initial data that approach zero sufficiently rapidly as $t \to \pm\infty$. The solution of the KdV equation was found as the external potential of a scattering problem. The method is now referred to as IST in the mathematical and physical community, while engineers prefer the term NFT.

In 1968 Lax [46] gave a mathematical formalization of the method, by relating the nonlinear evolution equation to a pair of operators with invariant eigenvalues—now referred to as Lax pair. At that time a Lax pair was know only for the KdV equation and, thus, the method could not be applied to other nonlinear evolution equations, despite being formalized in a general manner.

In 1972 Zakharov and Shabat [10] demonstrated that the normalized NLSE

$$j\frac{\partial q}{\partial z} = \frac{\partial^2 q}{\partial t^2} + 2\sigma||q||^2 q, \tag{3.2}$$

for the signal $q = q(z,t)$ can be solved with the IST method and provided a Lax pair associated with the NLSE.

After 1972, the Lax pairs associated with numerous nonlinear evolution equations have been found. Importantly, in 1974 Manakov [33] derived a Lax pair for the normalized ME

$$\begin{cases} j\dfrac{\partial q_1}{\partial z} = \dfrac{\partial^2 q_1}{\partial t^2} + 2\sigma(|q_1| + |q_2|)^2 q \\ j\dfrac{\partial q_2}{\partial z} = \dfrac{\partial^2 q_2}{\partial t^2} + 2\sigma(|q_1| + |q_2|)^2 q_2 \end{cases} . \tag{3.3}$$

3.1.2 Lax approach for the nonlinear Fourier transform

A Lax pair is defined [11] as a pair of operators (L, M) depending on the space coordinate z that satisfies the compatibility equation

$$\frac{\mathrm{d}L}{\mathrm{d}z} = ML - LM. \tag{3.4}$$

[1] In this thesis, spatial and temporal variables are exchanged with respect to the original notation, to be consistent with the notation used in optical fiber propagation.

Equation (3.4) induces a nonlinear evolution equation

$$\frac{\partial q}{\partial z} = K(q) \tag{3.5}$$

for $q = q(z,t)$, where K is an nonlinear operator that depends on the potential q, and its derivatives with respect to the time t, i.e., $K = K(q, |q^2|, q_t, q_{tt}, \dots)$.

Example 1. The KdV equation (3.1) is associated to the Lax pair $L = D + q$ and $M = [3D^3 + Dq + qD]/2$ where $D = \frac{\partial}{\partial t}$.

Proof. With some calculations one easily verifies that the compatibility equation (3.4) induces Eq. (3.1). □

Example 2. The Lax pair associated with the scalar NLSE is

$$L = j \begin{pmatrix} D & -q \\ -\sigma q^* & -D \end{pmatrix} \tag{3.6}$$

and

$$M = \begin{pmatrix} 2j\lambda^2 - j\sigma|q|^2 & -2\lambda q - jq_t \\ 2\lambda\sigma q^* - j\sigma q_t^* & -2j\lambda^2 + j\sigma|q|^2 \end{pmatrix} \tag{3.7}$$

with D being the derivative operator $D = \frac{\partial}{\partial t}$.

Proof. The compatibility equation $L_z = ML - LM$ holds four equations, one for each matrix component. Here, consider the equation $(L_z)_{12} = (ML)_{12} - (LM)_{12}$. Performing the calculations one obtains that

$$\begin{aligned} -jq_z = (L_z)_{12} &= (ML)_{12} - (LM)_{12} \\ &= j\left\{-4j\lambda^2 q + 2j\sigma|q|^2 q + [2\lambda q + jq_t]\, D + 2\lambda q_t + jq_{tt}\right\}; \end{aligned} \tag{3.8}$$

which implies that

$$\begin{cases} -jq_z = j\left\{-4j\lambda^2 q + 2j\sigma|q|^2 q + 2\lambda q_t + jq_{tt}\right\} \\ j\,[2\lambda q + jq_t]\, D = 0 \end{cases}, \tag{3.9}$$

yielding

$$jq_z = q_{tt} + 2\sigma|q|^2 q. \tag{3.10}$$

□

Importantly, the eigenvalues of L, i.e., the $\lambda \in \mathbb{C}$ such that

$$L\mathbf{v} = \lambda\mathbf{v}, \tag{3.11}$$

are independent of z, as a consequence of Eq. (3.4) [11]. On the other hand, considering Eq. (3.4) and the derivative with respect to z of Eq. (3.11), one obtains that the eigenvector $\mathbf{v}$ evolves in z according to [11]

$$\mathbf{v}_z = M\mathbf{v}. \tag{3.12}$$

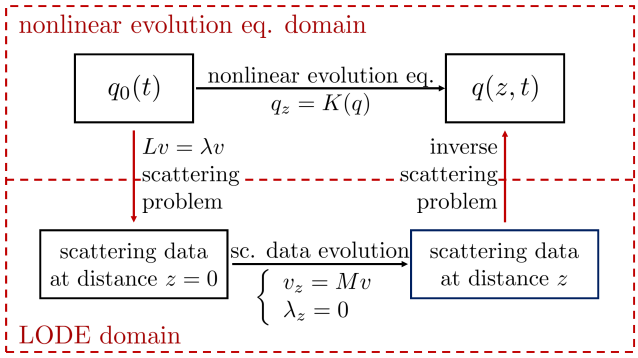


Figure 3.1: Lax approach for the solution of an IVP problem.

The Lax approach can be used to solve an IVP for a given nonlinear evolution equation, as sketched in Fig. 3.1, considering an associated scattering problem [11, 47]. Given the IVP

$$\begin{cases} q(0,t) = q_0(t) & \text{initial value} \\ q_z = K(q) & \text{evolution equation} \end{cases}, \tag{3.13}$$

the solution $q(z,t)$ can be obtained with the following approach.

1. Find, if it exists, a Lax pair (L, M) associated to the evolution equation, i.e., whose compatibility equation Eq. (3.4) induces the evolution equation.

2. For $z = 0$, solve the scattering problem, $L\mathbf{v} = \lambda\mathbf{v}$, where L contains $q_0(t)$ as a potential, and find the scattering data at $z = 0$. This is the direct scattering problem.

3. Find the evolution of the scattering data at a given distance z, as follows from Eq. (3.12).

4. Recover the potential $q(z,t)$ from the scattering data at distance z. This is the inverse scattering problem.

3.2 The nonlinear Fourier transform for the localized vector nonlinear Schrödinger equation

This section reviews the basic principles of the NFT for the general M-dimensional VNLSE with vanishing boundary conditions. The normalized VNLSE is

$$j\frac{\partial \mathbf{q}}{\partial z} = \frac{\partial^2 \mathbf{q}}{\partial t^2} + 2\sigma||\mathbf{q}||^2\mathbf{q}, \tag{3.14}$$

where $\mathbf{q} = \mathbf{q}(z,t)$ is an M dimensional vector of complex valued functions belonging to $L^1(\mathbb{R})$, meaning that $\mathbf{q}(z,t)$ vanishes rapidly as $t \to \pm\infty$ [8, 9, 48]. The parameter σ is equal to ± 1: the case with $\sigma = +1$ is referred to as focusing or antisymmetric case, while $\sigma = -1$ is the defocusing or symmetric case.

The Lax pair for Eq. (3.14) is the pair of operators of dimension $(M+1) \times (M+1)$

$$L = j\begin{pmatrix} \frac{\partial}{\partial t} & -\mathbf{q} \\ -\sigma\mathbf{q}^\dagger & -\frac{\partial}{\partial t}\mathbf{I}_M \end{pmatrix} \tag{3.15}$$

and

$$M = \begin{pmatrix} 2j\lambda^2 - j\sigma\mathbf{q}\mathbf{q}^\dagger & -2\lambda\mathbf{q} - j\mathbf{q}_t \\ 2\lambda\sigma\mathbf{q}^\dagger - j\sigma\mathbf{q}_t^\dagger & -2j\lambda^2\mathbf{I}_M + j\sigma\mathbf{q}^\dagger\mathbf{q} \end{pmatrix} \tag{3.16}$$

as can be proved following the same procedure as in Example 2.

In the following subsections, unless otherwise specified, the dependence on the space variable z is dropped. Indeed, both the FNFT—the direct scattering problem—and the BNFT—the inverse scattering problem—operations are independent of z. Nevertheless, one should remember that the quantities and functions involved depend also on z, e.g., the nonlinear spectrum $\boldsymbol{\rho}(\lambda) = \boldsymbol{\rho}(z,\lambda)$, and the eigenvector $\boldsymbol{\phi}(t,\lambda) = \boldsymbol{\phi}(z,t,\lambda)$.

The notations used in this Chapter (and in this thesis) are based on those used in the Chapter 4 "Matrix nonlinear Schrödinger systems" of [9] considering the special case with, in the notation of [9], $N = 1$, $M \geq 1$ being the dimension of the VNLSE, $\mathbf{Q} = \mathbf{q}$, $\mathbf{R} = -\sigma\mathbf{Q}^\dagger = -\sigma\mathbf{q}^\dagger$, and exchanging the spatial and temporal variables as $x \to t$ and $t \to z$.

3.2.1 Normalization procedure

Since the NFT can be applied to the normalized VNLSE, this subsection briefly describes a normalization procedure to obtain Eq. (3.14) from the dimensional VNLSE Eq. (2.10), which, we recall, describes the propagation of the optical signal in an optical fiber, under some assumptions (see Sec. 2 for more details).

Consider the three normalization parameters

$$\begin{cases} P_0 & [\mathrm{W}] \\ Z_0 = 2/(\gamma P_0) & [\mathrm{m}] \\ T_0 = \sqrt{|\beta_2|Z_0/2} & [\mathrm{s}] \end{cases} \tag{3.17}$$

where P_0 is a free parameter. Equation (2.10) can be reduced to the normalized form (3.2) considering

$$\begin{array}{l} \sigma = -\mathrm{sgn}(\beta_2) \\ \boldsymbol{q} \to \sqrt{\kappa P_0}\boldsymbol{q} \\ t \to t/T_0 \\ z \to -\sigma z/Z_0 \end{array}, \tag{3.18}$$

where κ accounts for the nonlinear coupling between spatial modes, and $\kappa = 1$ when $M = 1$ and $\kappa = 8/9$ for $M = 2$ (See Section 2.2). The case with $\sigma = 1$, obtained for $\beta_2 < 0$, corresponds to the anomalous dispersion case, typical of SMF.

In the following, unless otherwise stated, we will indicate with L the channel length in meters, and with $\mathcal{L}$ the normalized channel length, i.e., $\mathcal{L} = -\sigma L/Z_0$.

Remark 3. There exist some different notations for the NFT. The notations considered in this work are similar to [8, 9, 11, 49]. Different notations do not affect the concept of the NFT, but might slightly change some properties with a different sign or a 2 factor. For example, the authors of [18, 50] consider a slightly different normalized NLSE, and, in that case, the propagation of the nonlinear spectrum is described by the factor $e^{-2j\lambda^2\mathcal{L}}$, rather than $e^{-4j\lambda^2\mathcal{L}}$ (see Proposition 6).

3.2.2 Forward nonlinear Fourier transform

The FNFT operation consists in obtaining the nonlinear spectrum from the normalized M-dimensional optical signal $\mathbf{q}(t)$. The nonlinear spectrum, of dimension M, is composed of a continuous and dispersive part $\boldsymbol{\rho}(\lambda)$, defined on the nonlinear frequencies $\lambda \in \mathbb{R}$, and some discrete components $\{\mathbf{C}_i\}_{i=1}^{N_{\mathrm{ds}}}$ defined for $\lambda_i \in \mathbb{C}^+$, corresponding to solitons. The FNFT consists in solving the $M+1$-component Zakharov-Shabat (Z-S) problem [9, 33], which is the first equation defining the Lax pair for the VNLSE Eq. (3.11).

The eigenvalue problem for the VNLSE $L\mathbf{v} = \lambda\mathbf{v}$ (see Eqs. (3.15)-(3.16)) can be written as $\mathbf{v}_t = P\mathbf{v}$, where $\mathbf{v} \in \mathbb{C}^{M+1\times 1}$ is an auxiliary $M+1$-dimensional function, and [9]

$$P = \begin{pmatrix} -j\lambda & \mathbf{q}(t) \\ -\sigma\mathbf{q}^{\dagger}(t) & j\lambda\mathbf{I}_M \end{pmatrix} = \begin{pmatrix} -j\lambda & q_1(t) & \dots & q_M(t) \\ -\sigma q_1^*(t) & j\lambda & & \\ \vdots & & \ddots & \\ -\sigma q_M^*(t) & & & j\lambda \end{pmatrix} \tag{3.19}$$

an $M+1 \times M+1$ coupling matrix containing the signal $\mathbf{q}(t)$ as an effective potential. The solutions of $\mathbf{v}_t = P\mathbf{v}$ fixed by the boundary conditions at either the trailing or leading end of the multidimensional pulse have the basis $\{\boldsymbol{\phi}, \bar{\boldsymbol{\phi}}\}$ and $\{\boldsymbol{\psi}, \bar{\boldsymbol{\psi}}\}$ with [9]

$$\overbrace{\boldsymbol{\phi}(t,\lambda) \sim \begin{pmatrix} 1 \\ \mathbf{0}_{M\times 1} \end{pmatrix} e^{-j\lambda t}}^{M+1\times 1}, \quad \overbrace{\bar{\boldsymbol{\phi}}(t,\lambda) \sim \begin{pmatrix} \mathbf{0}_{1\times M} \\ \mathbf{I}_M \end{pmatrix} e^{j\lambda t}}^{M+1\times M} \quad \text{as } t \to -\infty, \tag{3.20}$$

$$\underbrace{\boldsymbol{\psi}(t,\lambda) \sim \begin{pmatrix} \mathbf{0}_{1\times M} \\ \mathbf{I}_M \end{pmatrix} e^{j\lambda t}}_{M+1\times M}, \quad \underbrace{\bar{\boldsymbol{\psi}}(t,\lambda) \sim \begin{pmatrix} 1 \\ \mathbf{0}_{M\times 1} \end{pmatrix} e^{-j\lambda t}}_{M+1\times 1} \quad \text{as } t \to +\infty. \tag{3.21}$$

The latter $\{\boldsymbol{\phi}, \bar{\boldsymbol{\phi}}, \boldsymbol{\psi}, \bar{\boldsymbol{\psi}}\}$ are referred to as canonical solutions, and are related to the Jost functions [11]. Consequently, there exist $a(\lambda) \in \mathbb{C}$, $\bar{\mathbf{a}}(\lambda) \in \mathbb{C}^{M\times M}$, $\mathbf{b}(\lambda) \in \mathbb{C}^{M\times 1}$, and $\bar{\mathbf{b}}(\lambda) \in \mathbb{C}^{1\times M}$—referred to as scattering data—such that

$$\boldsymbol{\phi}(t,\lambda) = \boldsymbol{\psi}(t,\lambda)\mathbf{b}(\lambda) + \bar{\boldsymbol{\psi}}(t,\lambda)a(\lambda), \tag{3.22}$$

$$\bar{\boldsymbol{\phi}}(t,\lambda) = \boldsymbol{\psi}(t,\lambda)\bar{\mathbf{a}}(\lambda) + \bar{\boldsymbol{\psi}}(t,\lambda)\bar{\mathbf{b}}(\lambda). \tag{3.23}$$

The scattering data are defined on the $\lambda \in \mathbb{C}$ such that Eq. (3.22) and (3.23) are well defined.

The scattering coefficients can be obtained through the evaluation of the solution $\boldsymbol{\phi}(t,\lambda)$, defined by the boundary condition at $-\infty$, at the opposite end of the interval as

$$a(\lambda) = \lim_{t\to+\infty} \phi_1(t,\lambda)e^{+j\lambda t}, \qquad b_m(\lambda) = \lim_{t\to+\infty} \phi_{m+1}(t,\lambda)e^{-j\lambda t}, \tag{3.24}$$

for $m = 1,\dots,M$, as follows from Eqs. (3.20-3.23).

Continuous spectrum If $\lambda \in \mathbb{R}$, the scattering data are well defined. The scattering matrix, defined as the $M+1 \times M+1$ matrix

$$\mathbf{S} = \begin{pmatrix} a(\lambda) & \bar{\mathbf{b}}(\lambda) \\ \mathbf{b}(\lambda) & \bar{\mathbf{a}}(\lambda) \end{pmatrix}, \tag{3.25}$$

is unimodular, i.e., $\det \mathbf{S} = 1$ [8, 9]. The continuous nonlinear spectrum—also referred to as reflection coefficient—is eventually defined as the M-component vector

$$\boldsymbol{\rho}(\lambda) = \mathbf{b}(\lambda)/a(\lambda), \tag{3.26}$$

where we are assuming that $a(\lambda) \neq 0$ for $\lambda \in \mathbb{R}$.

Discrete spectrum If $\Im(\lambda) > 0$, from (3.22) and the asymptotics (3.20) and (3.21) it follows that $a(\lambda) = 0$. Moreover, since $a(\lambda)$ is analytic in $\mathbb{C}^+$ [8, 9] it follows that the number of zeros of $a(\lambda)$ in $\mathbb{C}^+$ is finite. The discrete nonlinear spectrum is the set of N_{DS} eigenvalues $\{\lambda_i\}_{i=1}^{N_{\mathrm{DS}}}$ such that

$$\lambda_i \in \mathbb{C}^+ \text{ and } a(\lambda_i) = 0 \tag{3.27}$$

and of the corresponding discrete spectral components, also referred to as norming constants, given by

$$\mathbf{C}_i = \mathbf{b}(\lambda_i)/a'(\lambda_i), \tag{3.28}$$

where $a'(\lambda) = \partial a/\partial\lambda$ and the zeros of $a(\lambda)$ are simple [2]. Discrete components correspond to solitons, particular solutions that maintain their shape (or return periodically to it) during propagation thanks to the balancing of nonlinear and dispersive effects.

[2]For the sake of simplicity, we assume that λ_i is a simple zero of $a(\lambda)$. We refer to [51] for an interesting work about communication using eigenvalues with higher multiplicity.

3.2.3 Backward nonlinear Fourier transform

The BNFT is the inverse operation to retrieve the time domain signal from the nonlinear spectrum. Several methods can be used to perform this operation. While considering the BNFT as a Riemann-Hilbert problem [52] might be the most general method, we found more convenient (for the purposes of this thesis) to consider Gelfand-Levitan-Marchenko equation (GLME)-based methods. Moreover, the Darboux transform (DT) is useful when discrete eigenvalues are involved, but requires additional methods when the continuous spectrum is not zero [53].

The M-dimensional GLME, i.e., the vector GLME (VGLME), associated with Eq. (3.14) is the integral equation

$$\mathbf{K}(x,y) - \sigma\mathbf{F}^{\dagger}(x+y) + \sigma \int_{x}^{\infty}\int_{x}^{\infty} \mathbf{K}(x,r)\mathbf{F}(r+s)\mathbf{F}^{\dagger}(s+y)\,dsdr = 0\,, \quad (3.29)$$

for the unknown M-component function $\mathbf{K}(x,y)$ [9], as derived in Section A.3. The M-dimensional kernel function $\mathbf{F}(x)$ depends on the nonlinear spectrum and is defined as

$$\mathbf{F}(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \boldsymbol{\rho}(\lambda)e^{j\lambda x}\,d\lambda - j\sum_{i=1}^{N}\mathbf{C}_i e^{j\lambda_i x}. \quad (3.30)$$

The time domain signal is obtained solving the GLME for $\mathbf{K}(x,y)$ as $\mathbf{q}(t) = -2\mathbf{K}(t,t)$ [9].

Note that the first addend in Eq. (3.30) is just an FT so that (3.30) can be written as

$$\mathbf{F}(x) = \frac{1}{2\pi}\mathcal{F}\left\{\boldsymbol{\rho}(\lambda)\right\}(-x/(2\pi)) - j\sum_{i=1}^{N}\mathbf{C}_i e^{j\lambda_i x}. \quad (3.31)$$

3.2.4 Properties and remarks

Several interesting properties of the NFT for the VNLSE can be found in [9], and even more for the scalar NLSE can be found in [8, 11]. This subsection reports some useful (for this thesis) properties together with others that, to the best of my knowledge, cannot be found in the literature (at least, for the M-dimensional VNLSE). In particular, Propositions 10 and 11, demonstrated here for $M \geq 1$, have been originally proved in our works: Proposition 10 in [54] for $M = 1$ and Proposition 11 in [40] for $M = 2$.

Proposition 4. *In the defocusing regime, i.e., when $\sigma = -1$, the discrete spectrum is empty.*

Proof. When $\sigma = -1$, the operator L is Hermitian, i.e., $LL^{\dagger} = \mathbf{I}_{M+1}$. As a consequence, all the eigenvalues are real, and, thus, there do not exist imaginary eigenvalues. □

Proposition 5. *When $||\mathbf{q}||_1 \ll 1$, the discrete spectrum is absent and the continuous nonlinear spectrum tends to the linear spectrum as*

$$\boldsymbol{\rho}(\lambda) \to -\sigma \mathbf{Q}^\dagger(-\lambda/\pi) = -\sigma \int_{-\infty}^{+\infty} \mathbf{q}^\dagger(t) e^{-j2\lambda t}\, dt, \tag{3.32}$$

with $\mathbf{Q}(f)$ being the FT of $\mathbf{q}(t)$.

Proof. The Proposition is proved in [11] for $M = 1$. For $M \geq 1$, it follows from the definition of the kernel function Eq. (3.30) and the VGLME Eq. (3.29) in the asymptotics when $||\mathbf{q}||_1$ is small. Indeed,

$$\begin{array}{rcl} \int_{-\infty}^{+\infty} \mathbf{Q}(f) e^{j2\pi ft}\, df = \mathbf{q}(t) = & -2\mathbf{K}(t,t) \to -2\sigma \mathbf{F}^\dagger(2t) \\ = & -\sigma/\pi \int_{-\infty}^{+\infty} \boldsymbol{\rho}^\dagger(\lambda) e^{-2j\lambda t}\, d\lambda. \end{array} \tag{3.33}$$

□

Proposition 6. *The propagation of the nonlinear spectrum to normalized distance $\mathcal{L}$ is equivalent to the multiplication of each NFT spectral component by $e^{-4j\lambda^2\mathcal{L}}$—sometimes referred to as channel filter. The discrete eigenvalues are constant in space i.e., $\lambda_i(z) \equiv \lambda_i$.*

Proof. The proof of the first part follows as in [11]. The propagation equation for the eigenvector $\mathbf{v}$ is $\mathbf{v}_z = M\mathbf{v}$, where M is given by Eq. (3.16). When $t \to \pm\infty$, the operator M tends to

$$M_{\pm\infty} = \begin{pmatrix} 2j\lambda^2 & \mathbf{0}_{1\times M} \\ \mathbf{0}_{M\times 1} & -2j\lambda^2 \mathbf{I}_M \end{pmatrix}. \tag{3.34}$$

The canonical solutions $\boldsymbol{\phi}$, $\bar{\boldsymbol{\phi}}$, $\boldsymbol{\psi}$ and $\bar{\boldsymbol{\psi}}$ are defined in the Eqs. (3.20)-(3.21) for $z = 0$, and, therefore,

$$\begin{array}{l} \lim_{t\to\pm\infty} \phi_1(z,t,\lambda) = \lim_{t\to\pm\infty} \phi_1(0,t,\lambda) e^{2j\lambda^2 z} \\ \lim_{t\to\pm\infty} \phi_{m+1}(z,t,\lambda) = \lim_{t\to\pm\infty} \phi_{m+1}(0,t,\lambda) e^{-2j\lambda^2 z} \quad \text{for } m = 1,\ldots,M \end{array}, \tag{3.35}$$

where the dependence on z is now explicit. Let $\boldsymbol{\theta}(\mathcal{L},t,\lambda) = \boldsymbol{\phi}(\mathcal{L},t,\lambda) e^{-2j\lambda^2\mathcal{L}}$, then

$$\lim_{t\to-\infty} \boldsymbol{\theta}(\mathcal{L},t,\lambda) = \lim_{t\to-\infty} \boldsymbol{\phi}(\mathcal{L},t,\lambda) e^{-2j\lambda^2\mathcal{L}} = \lim_{t\to-\infty} \begin{pmatrix} 1 \\ \mathbf{0}_{M\times 1} \end{pmatrix} e^{-j\lambda t}, \tag{3.36}$$

as follows from Eqs. (3.20)-(3.35). Consequently, in $z = \mathcal{L}$ $\boldsymbol{\theta}$ has the same boundary condition at $t \to -\infty$ of $\boldsymbol{\phi}$ in $z = 0$. Thus, the scattering data in $z = \mathcal{L}$ can be defined through the value of $\boldsymbol{\theta}(\mathcal{L},t,\lambda)$ for $t \to +\infty$, i.e.,

$$\begin{array}{rl} a(\lambda,\mathcal{L}) & = \lim_{t\to+\infty} \theta_1(\mathcal{L},t,\lambda) e^{+j\lambda t} = \lim_{t\to+\infty} \phi_1(\mathcal{L},t,\lambda) e^{-2j\lambda^2\mathcal{L}} e^{+j\lambda t} \\ & = \lim_{t\to+\infty} \phi_1(0,t,\lambda) e^{+j\lambda t} = a(0,\lambda), \end{array} \tag{3.37}$$

$$\begin{aligned} b_m(\lambda,\mathcal{L}) &= \lim_{t\to+\infty}\theta_{m+1}(\mathcal{L},t,\lambda)e^{-j\lambda t} = \lim_{t\to+\infty}\phi_{m+1}(\mathcal{L},t,\lambda)e^{-2j\lambda^2\mathcal{L}}e^{-j\lambda t} \\ &= \lim_{t\to+\infty}\phi_{m+1}(0,t,\lambda)e^{-4j\lambda^2\mathcal{L}}e^{-j\lambda t} = b_m(0,\lambda)e^{-4j\lambda^2\mathcal{L}}. \end{aligned} \tag{3.38}$$

Finally,

$$\boldsymbol{\rho}(\mathcal{L},\lambda) = \mathbf{b}(\mathcal{L},\lambda)/a(\mathcal{L},\lambda) = \mathbf{b}(0,\lambda)/a(0,\lambda)e^{-4j\lambda^2\mathcal{L}} = \boldsymbol{\rho}(0,\lambda)e^{-4j\lambda^2\mathcal{L}}. \tag{3.39}$$

The second part of the proposition comes directly from the definition of Lax pair. Indeed, the compatibility equation (3.4) holds if an only if the eigenvalues of $L(z)$ do not depend on z (i.e., L is isospectral) [11], or, equivalently from Eq. (3.37). □

Proposition 7. *For any* $\lambda \in \mathbb{R}$

$$|a(\lambda)|^2 + \sigma\mathbf{b}(\lambda)^\dagger\mathbf{b}(\lambda) = |a(\lambda)|^2 + \sigma\sum_{m=1}^{M}|b_m(\lambda)|^2 = 1. \tag{3.40}$$

Proof. The property follows from the asymptotic behavior of $(\boldsymbol{\sigma}\boldsymbol{\phi})^\dagger\boldsymbol{\phi}$, $(\boldsymbol{\sigma}\boldsymbol{\psi})^\dagger\boldsymbol{\psi}$, and $\left(\boldsymbol{\sigma}\bar{\boldsymbol{\psi}}\right)^\dagger\bar{\boldsymbol{\psi}}$, which remains constant in time [9]. □

Proposition 8. *The nonlinear analog of Parseval's identity that relates the energy of time domain signal to the energy defined through the nonlinear spectrum, is*

$$\int_{-\infty}^{+\infty}\sum_{m=1}^{M}|q_m(t)|^2\,dt = 4\sum_{k=1}^{N}\Im\{\lambda_k\} + \frac{\sigma}{\pi}\int_{-\infty}^{+\infty}\log\left(1+\sigma\sum_{m=1}^{M}|\rho_m(\lambda)|^2\right)d\lambda. \tag{3.41}$$

Proof. The relation follows from the trace formula [8, 9]

$$\int_{-\infty}^{+\infty}\sum_{m=1}^{M}|q_m(t)|^2\,dt = 4\sum_{k=1}^{N}\Im\{\lambda_k\} - \frac{\sigma}{\pi}\int_{-\infty}^{+\infty}\log\left(|a(\lambda)|^2\right)d\lambda. \tag{3.42}$$

and using that $|a(\lambda)|^2 = (1+\sum_{m=1}^{M}|\rho_m(\lambda)|^2)^{-1}$ as follows from Prop. 7. □

Proposition 9. *If* $\sigma = 1$ *and for* $\lambda \in \mathbb{R}$*, the modulus of the scattering data is upper limited with*

$$\begin{aligned} &|a(\lambda)| \le 1 \\ &|b_m(\lambda)| \le 1 \quad \textit{for } m = 1,\dots,M \end{aligned}. \tag{3.43}$$

If $\sigma = -1$ *and for* $\lambda \in \mathbb{R}$*, the modulus of the nonlinear spectrum is upper limited with*

$$|\rho_m(\lambda)| \le 1 \quad \textit{for } m = 1,\dots,M\ . \tag{3.44}$$

Proof. The relations follow from Prop. 7. Indeed for $\sigma = 1$

$$0 \leq |a(\lambda)|^2 = 1 - \underbrace{\sum_{m=1}^{M} |b_m(\lambda)|^2}_{\geq 0} \leq 1, \tag{3.45}$$

and

$$0 \leq b_m = 1 - \underbrace{|a(\lambda)|^2}_{\geq 0} - \underbrace{\sum_{m \neq k=1}^{M} |b_k(\lambda)|^2}_{\geq 0} \leq 1. \tag{3.46}$$

Similarly, for $\sigma = -1$

$$1 - \sum_{m=1}^{M} |b_m(\lambda)|^2 / |a(\lambda)|^2 = (|a(\lambda)|^2)^{-1}, \tag{3.47}$$

and, thus,

$$0 \leq |\rho_m(\lambda)|^2 = 1 - \underbrace{(|a(\lambda)|^2)^{-1}}_{\geq 0} - \underbrace{\sum_{m \neq k=1}^{M} |\rho_k(\lambda)|^2}_{\geq 0} \leq 1, \tag{3.48}$$

□

Proposition 10. (*NFT* Causality Property) *Given a generic time instant τ, for $t \geq \tau$ the time domain signal $\mathbf{q}(t)$ depends only on the values of $\mathbf{F}(y)$ for $y \geq 2\tau$.*

Proof. The proof follows as in [54] for the scalar NLSE. Since $\mathbf{q}(t) = -2\mathbf{K}(t,t)$, one should obtain the solution $\mathbf{K}(t,t)$ of Eq. (3.29) for all $t \geq \tau$. Given $t_1, t_2 \geq \tau$, the solution $\mathbf{K}(t_1,t_2) = \sigma \mathbf{F}^\dagger(t_1+t_2) - \sigma \int_{t_1}^{\infty}\int_{t_1}^{\infty} \mathbf{K}(t_1,r)\mathbf{F}(r+s)\mathbf{F}^\dagger(s+t_2)\,\mathrm{d}r\mathrm{d}s$ depends on $\mathbf{F}(y)$ for all $y \geq 2\tau$ (since $r+s \geq 2t_1 \geq 2\tau$ and $s+t_2 \geq t_1+t_2 \geq 2\tau$) and on all the solutions $\mathbf{K}(t_1,r)$ for $r \geq t_1$. Nevertheless, since $t_1 \geq \tau$, $\mathbf{K}(t_1,r)$ is one of the solutions $\mathbf{K}(t_1,t_2)$ for $t_1, t_2 \geq \tau$. Therefore, for $t_1, t_2 \geq \tau$, $\mathbf{K}(t_1,t_2)$ depends only on $\mathbf{F}(y)$ for $y \geq 2\tau$. Consequently, considering $t_1 = t_2 = t$, one obtains that, for $t \geq \tau$, $q(t) = -2\mathbf{K}(t,t)$ depends only on $\mathbf{F}(y)$ for $y \geq 2\tau$. □

Proposition 11. *If R is an unitary $M \times M$ matrix, i.e., $R^\dagger R = \mathbf{I}_M$. Then,*

$$FNFT\left(\mathbf{q}R^{\mathrm{T}}\right) = R^* FNFT(\mathbf{q})\,. \tag{3.49}$$

Importantly, for the internal coherence of the notation used in this thesis [9] $\mathbf{q}$ is a row vector, while the nonlinear spectrum (i.e., the result of the FNFT) is a column vector. This is the reason why the argument of the FNFT in Eq. (3.49) is $\mathbf{q}R^{\mathrm{T}}$, which is equal to $(R\mathbf{q}^T)^{\mathrm{T}}$.

In the single-polarization case, Eq. (3.49) reduces to [12]

$$FNFT\left[e^{j\phi}q\right] = e^{-j\phi}FNFT[q] = e^{-j\phi}\rho, \tag{3.50}$$

while for $M = 2$ it becomes

$$FNFT\left[\begin{pmatrix} R_{11}q_1 + R_{12}q_2 \\ R_{21}q_1 + R_{22}q_2 \end{pmatrix}^{\mathrm{T}}\right] = \begin{pmatrix} R_{11}^*\rho_1 + R_{12}^*\rho_2 \\ R_{21}^*\rho_1 + R_{22}^*\rho_2 \end{pmatrix}, \tag{3.51}$$

with $(\rho_1, \rho_2)^T = FNFT[(q_1, q_2)]$.

Proof. Firstly, let us prove that if $\mathbf{v}$ is a solution of the Z-S problem $\mathbf{v}_t = P(\mathbf{q})\mathbf{v}$, where $P(\mathbf{q})$ indicates that P is associated with the potential $\mathbf{q}(t)$, then

$$\mathbf{u} \doteqdot \mathcal{R}\mathbf{v} = \begin{pmatrix} 1 & \mathbf{0}_{1\times M} \\ \mathbf{0}_{M\times 1} & R^* \end{pmatrix}\mathbf{v} = \begin{pmatrix} v_1 \\ R^*\bar{\mathbf{v}} \end{pmatrix}, \tag{3.52}$$

where $\bar{\mathbf{v}} = (v_2, \ldots, v_{M+1})^{\mathrm{T}}$ is a solution of the Z-S problem associated with $\mathbf{q}R^{\mathrm{T}}$, i.e., $\mathbf{u}_t = P(\mathbf{q}R^{\mathrm{T}})\mathbf{u}$. This holds if and only if

$$\begin{pmatrix} v_{1,t} \\ R^*\bar{\mathbf{v}}_t \end{pmatrix} = \begin{pmatrix} -j\lambda & \mathbf{q}R^{\mathrm{T}} \\ -\sigma(\mathbf{q}R^{\mathrm{T}})^\dagger & j\lambda\mathbf{I}_M \end{pmatrix}\begin{pmatrix} v_1 \\ R^*\bar{\mathbf{v}} \end{pmatrix} = \begin{pmatrix} -j\lambda v_1 + \mathbf{q}R^{\mathrm{T}}R^*\bar{\mathbf{v}} \\ -\sigma R^*\mathbf{q}^\dagger v_1 + j\lambda R^*\bar{\mathbf{v}} \end{pmatrix}. \tag{3.53}$$

Using that $RR^\dagger$ implies that $\mathbf{I}_M = R^\dagger R = R^{\mathrm{T}}R^*$ implies that the first equation can be rewritten as

$$v_{1,t} = -j\lambda v_1 + \mathbf{q}\mathbf{I}_M\bar{\mathbf{v}} = -j\lambda v_1 + \sum_{m=1}^{M} q_m v_{m+1} \tag{3.54}$$

which is exactly the first equation induced by $\mathbf{v}_t = P(\mathbf{q})\mathbf{v}$, our hypothesis. The second equation (of dimension M) can be rewritten as

$$R^*\bar{\mathbf{v}}_t = -\sigma R^*\mathbf{q}^\dagger v_1 + j\lambda R^*\bar{\mathbf{v}} \tag{3.55}$$

which is equivalent to the second (matrix) equation induced by $\mathbf{v}_t = P(\mathbf{q})\mathbf{v}$ (just multiply Eq. (3.55) by R^{T}).

Secondly, $\mathcal{R}\boldsymbol{\phi}$ has the same boundary condition (3.20) as $\boldsymbol{\phi}$ at $-\infty$ and $\mathcal{R}\boldsymbol{\phi}$ is a solution of the Z-S problem $\mathbf{u}_t = P(\mathbf{q}R^{\mathrm{T}})\mathbf{u}$. Consequently, the nonlinear spectrum associated with $\mathbf{q}R^{\mathrm{T}}$ can be obtained from the values at $+\infty$ of $\mathcal{R}\boldsymbol{\phi}$, through Eq. (3.24). The scattering data $a_R(\lambda)$ associated with $\mathbf{q}R^{\mathrm{T}}$ is obtained through the first component of $\mathcal{R}\boldsymbol{\phi}$, which does not change and, therefore, $a_R(\lambda) = a(\lambda)$. On the other hand, the scattering data $\mathbf{b}_R(\lambda)$ associated with $\mathbf{q}R^{\mathrm{T}}$ is obtained through the following M components of $\mathcal{R}\boldsymbol{\phi}$, and, thus $\mathbf{b}_R(\lambda) = R^*\mathbf{b}(\lambda)$. Finally, the nonlinear spectrum changes as

$$\boldsymbol{\rho}_R(\lambda) = \mathbf{b}_R(\lambda)/a_R(\lambda) = R^*\mathbf{b}(\lambda)/a(\lambda) = R^*\boldsymbol{\rho}_R(\lambda). \tag{3.56}$$

□

3.3 The nonlinear Fourier transform for the scalar nonlinear Schrödinger equation

This section briefly resumes the concepts described in the previous section for the scalar NLSE. A wide literature about this topic exists [7, 8, 10, 11, 47].

The normalized scalar NLSE Eq. (3.2) is an integrable equation. This means that, given a certain initial value $q(0,t) = q_0(t)$, a solution $q(z,t) \in L^1(\mathbb{R})$ with vanishing boundary conditions (i.e., $|q(z,t)| \to 0$ as $t \to \pm\infty$) can be obtained through the NFT method. In this case the nonlinear spectrum is made of a continuous part

$$\rho(\lambda) = b(\lambda)/a(\lambda) \tag{3.57}$$

and N_{DS} discrete components

$$C_i = b(\lambda_i)/a_\lambda(\lambda_i), \quad a(\lambda_i) = 0, \quad \lambda_i \in \mathbb{C}^+. \tag{3.58}$$

The scattering data $a(\lambda)$ and $b(\lambda)$ are defined through the solutions $\boldsymbol{\nu}(t,\lambda) \in \mathbb{C}^{2\times 1}$ of the Z-S problem $L\mathbf{v} = \lambda\mathbf{v}$ where

$$L = j \begin{pmatrix} \partial/\partial t & -q \\ -\sigma q^* & -\partial/\partial t \end{pmatrix}. \tag{3.59}$$

The scattering data can be obtained as

$$a(\lambda) = \lim_{t\to+\infty} \phi_1(t,\lambda) e^{j\lambda t}, \tag{3.60}$$

$$b(\lambda) = \lim_{t\to+\infty} \phi_2(t,\lambda) e^{-j\lambda t} \tag{3.61}$$

where $\boldsymbol{\phi}(t,\lambda)$ is the solution of $L\boldsymbol{\nu} = \lambda\boldsymbol{\nu}$ with

$$\lim_{t\to-\infty} \boldsymbol{\phi}(t,\lambda) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-j\lambda t}. \tag{3.62}$$

The inverse operation to recover the time domain signal $q(t)$ from the nonlinear spectrum can be achieved via the solution of the GLME

$$K(x,y) - \sigma F^*(x+y) + \sigma \int_x^{+\infty} \int_x^{+\infty} K(x,r) F(r+s) F^*(s+y) ds dr = 0 \tag{3.63}$$

and letting $q(t) = -2K(t,t)$. The integral kernel $F(y)$ is defined as

$$F(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \rho(\lambda) e^{j\lambda x} d\lambda - j \sum_{i=1}^{N_{\mathrm{DS}}} C_i e^{j\lambda_i x}, \tag{3.64}$$

and it reduces to

$$F(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \rho(\lambda) e^{j\lambda x} d\lambda = \frac{1}{2\pi} \mathcal{F}\left\{\rho(\lambda)\right\}(-x/(2\pi)) \tag{3.65}$$

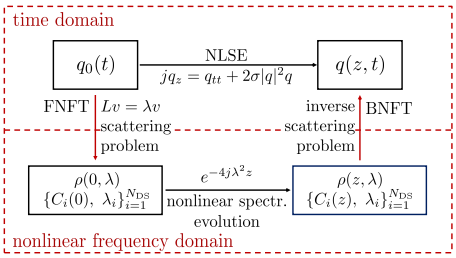


Figure 3.2: Lax approach for the solution of an IVP problem associated with the NLSE.

when the discrete spectrum is empty.

The evolution to normalized distance $\mathcal{L}$ of the nonlinear spectrum is given by multiplication for $e^{-4j\lambda^2\mathcal{L}}$, i.e.,

$$\rho(\mathcal{L},\lambda) = \rho(0,\lambda)e^{-4j\lambda^2\mathcal{L}} \tag{3.66}$$

and

$$C_i(\mathcal{L}) = C_i(0)e^{-4j\lambda^2\mathcal{L}} \quad \text{for } i = 1,\ldots,N_{\rm DS}. \tag{3.67}$$

Therefore, the solution $q = q(z,t)$ of the IVP

$$\begin{cases} j\frac{\partial q}{\partial z} = \frac{\partial^2 q}{\partial t^2} + 2\sigma||q||^2 q, & \text{NLSE} \\ q(0,t) = q_0(t) & \text{initial value} \end{cases} \tag{3.68}$$

can be obtained as sketched in Fig. 3.2 through:

1. Direct scattering problem: perform a FNFT to obtain the nonlinear spectrum $\rho(0,\lambda)$ and $\{C_i(0),\lambda_i\}_{i=1}^{N_{DS}}$ in $z = 0$, starting from the initial value $q_0(t)$;

2. Propagate the nonlinear spectrum according to Eqs. (3.66)-(3.67) to obtain the nonlinear spectrum $\rho(z,\lambda)$ and $\{C_i(z),\lambda_i\}_{i=1}^{N_{DS}}$ at distance z;

3. Inverse scattering problem: perform a BNFT to obtain the time domain signal $q(z,t)$ at distance z, starting from the nonlinear spectrum $\rho(z,\lambda)$ and $\{C_i(z),\lambda_i\}_{i=1}^{N_{DS}}$ at distance z .

3.3.1 Understanding the nonlinear Fourier transform

Due to its analogy with the standard linear FT, the IST method is more often referred to as NFT in the engineering community. Indeed, the NFT can be understood as the nonlinear analog of the standard FT, despite the fact that a compact form (similar to (2.11)) does not exist for the NFT.

The FT is a powerful tool to solve linear IVPs as it allows to describe the complicated evolution in space (given by a linear differential equation) in a simple way in the Fourier domain. For example, the solution of the linear IVP

$$\begin{cases} jq_z(z,t) = q_{tt}(z,t) \\ q(0,t) = q_0(t) \end{cases} \tag{3.69}$$

for $q = q(z,t)$ can be obtained straightforwardly in the Fourier domain. Indeed, performing an FT of the evolution equation—and using the notation $\hat{q}(f) = \mathcal{F}\{q(t)\}(f)$—one obtains that $j\hat{q}_z(z,f) = (2\pi jf)^2\hat{q}(z,f)$, implying that $\hat{q}(z,f) = e^{j4\pi^2 f^2 z}A$, with A a constant (in space z) term. The solution of the IVP is achieved as:

1. Obtain the Fourier spectrum $\hat{q}_0(f)$ of $q_0(t)$ performing an FT.

2. The evolution of the spectrum is obtained as $\hat{q}(z,f) = e^{j4\pi^2 f^2 z}\hat{q}_0(f)$.

3. Obtain the solution $q(z,t)$ of the IVP performing an inverse FT of $\hat{q}(z,f)$.

Unfortunately, the same can not be done if the equation associated with the IVP is nonlinear. Indeed, given

$$\begin{cases} jq_z(z,t) = q_{tt}(z,t) + 2\sigma|q(z,t|^2 q(z,t) \\ q(0,t) = q_0(t) \end{cases} \tag{3.70}$$

performing an FT of the evolution equation drives to

$$j\hat{q}_z(z,f) = (2\pi jf)^2\hat{q}(z,f) + 2\sigma\int_{-\infty}^{+\infty} |q(z,t|^2 q(z,t)e^{-j2\pi ft}\,dt. \tag{3.71}$$

The latter implies that there is not an *easy* formulation for the propagation of the spectrum.

Nevertheless, a similar scheme can be used considering the NFT instead of the FT, thanks to Eqs. (3.66)-(3.67). The three steps are:

1. Obtain the nonlinear Fourier spectrum $\rho(0,\lambda)$ and $\{C_i(0),\lambda_i\}_{i=1}^{N_{DS}}$ of $q_0(t)$ performing an FNFT—i.e., the direct NFT. This is the direct scattering.

2. The evolution of the nonlinear spectrum is obtained from Eqs. (3.66)-(3.67). Similar to the scalar case, the propagation of the spectrum is linear (in space), and it is the multiplication for an exponential factor.

3. Obtain the solution $q(z,t)$ of the IVP performing a BNFT—i.e., an inverse NFT—of the nonlinear spectrum.

3.3.2 The nonlinear Fourier transform in fiber Bragg gratings

The nonlinear Fourier transform related to the NLSE with vanishing boundary conditions in the normal dispersion regime $\sigma = -1$, describes not only the propagation of a signal in a lossless fiber with negative GVD parameter, but also the relation between a fiber Bragg grating and its reflection spectrum, where the waveform plays the role of the grating profile [55, 56]. Indeed, a fiber Bragg grating with high values of refractive index corresponds to an optical signal with high power, while a wide fiber Bragg grating corresponds to an optical signal with long duration [56].

When the optical signal has high power or long duration, and in this case its L^1-norm is large, the reflection spectrum $\rho(\lambda)$ is close to 1 (recall that $|\rho(\lambda)| \leq 1$ from Proposition 9) meaning that the light is almost totally reflected by the grating. In this case, it is well know that the inverse synthesis problem, i.e., the problem of reconstructing the fiber Bragg grating profile—equivalently, the waveform—from the reflected spectrum, is an ill-posed problem, for which numerous algorithms have been proposed [55–57].

In analogy with the numerical instabilities related to the synthesis of strongly reflective fiber Bragg gratings, in the following of this thesis, we will show similar numerical instabilities also in the anomalous dispersion—i.e., focusing—regime $\sigma = 1$.

3.3.3 The nonlinear Fourier transform with periodic boundary condition

As already mentioned at the beginning of this section, the NFT is a method to solve the IVP associated with a certain class of nonlinear equations, referred to as integrable equations. Despite the fact that this thesis is focused on the use of the NFT with vanishing boundary conditions, it is worth noting that the NFT for the NLSE can be defined also for signals with different boundary conditions. In particular, the application of the NFT with periodic boundary conditions to optical fiber communication has also gained attention in the past year [17, 52], as a way to overcome some of the issues related to the fulfillment of the vanishing boundary conditions, as will be clear in the following chapters.

In a nutshell, the nonlinear Fourier spectrum for periodic boundary conditions is made of a *main spectrum* $\mathbb{M}$, which corresponds to the discrete spectrum in the case with vanishing boundary conditions, and an *auxiliary spectrum* $\mathbb{A}$. The main spectrum is invariant during propagation. Unfortunately, the numerical computation of the BNFT is quite challenging [17, 52]. However, the area of the application of the NFT with periodic boundary conditions is slowly gaining attention [17, 52].

3.4 Numerical methods for the nonlinear Fourier transform for the localized vector nonlinear Schrödinger equation

This section presents some numerical methods for the NFT for the multidimensional VNLSE with vanishing boundary conditions, one for the FNFT and one for the BNFT operation. Moreover, this section derives an alternative method to compute the BNFT for $M = 2$ when both the discrete and the continuous spectrum are present, combining GLME-based methods with the DT. We remark that research about numerical methods for the NFT operations is still in progress, and various methods exist, in particular for the scalar NFT associated with the NLSE. Consequently, the aim of this section is not to provide optimal algorithms for the numerical NFTs but to give some general methods that can be used in various scenarios.

Regarding the scalar NFT, different methods are available for the FNFT and an interested reader can refer to [7, 11, 58–63, and reference therein]. On the other hand for the BNFT, GLME-based methods can deal with both discrete and continuous spectrum [41,50,58,64–66], while DT-based methods can be used for discrete eigenvalues (when the continuous spectrum is absent, or combining it with other methods) [7, 11, 14, 53, 67]; finally methods based on the inversion of the FNFT can be used when the discrete spectrum is absent [12].

As far as it concerns the ME and the VNLSE, some of the aforementioned methods have been extended to these cases: the DT for the ME [68], the Ablowitz–Ladik method for the FNFT for the ME and its inverse (when there is only the continuous spectrum) for the BNFT [49], and the trapezoidal integration method for the FNFT for the ME [68]. The extension to the multidimensional VNLSE of the Boffetta–Osborne method for the FNFT and of the Nyström-conjugate gradient (NCG) method for the BNFT (applicable to joint spectrum) have been derived in [40, 69], and are presented in the following sections.

3.4.1 Numerical forward nonlinear Fourier transform

This subsection presents a numerical method to recover the scattering data $a(\lambda)$ and $\mathbf{b}(\lambda)$ (an M-dimensional vector) starting from the time domain signal $\mathbf{q}(t)$, i.e. to solve the vector Z-S problem associated with the VNLSE [9]. The method considered here is a multidimensional extension of the Boffetta–Osborne method [58] (also known as the layer-peeling method [11]) developed for the scalar NLSE.

Let us assume that $|\mathbf{q}(t)| = 0$ for $|t| > T$ and consider a uniform grid with $t_n = -T + (n-1)\delta$ for $n = 1, \ldots, N_t + 1$, and discretization step $\delta = 2T/N_t$. Rather than solving the problem $\boldsymbol{\nu}_t = P\boldsymbol{\nu}$, the idea is to iteratively solve the

Cauchy problem,

$$\begin{cases} \boldsymbol{\phi}_t = P^{(n)}\boldsymbol{\phi} & \text{for } t \in (t_n - \delta/2, t_n + \delta/2], \\ \boldsymbol{\phi}(t_n - \delta/2) = \boldsymbol{\phi}^{(n)} & \text{initial value,} \end{cases} \tag{3.72}$$

to define the initial value for the following iteration as $\boldsymbol{\phi}^{(n+1)} = \boldsymbol{\phi}(t_n+\delta/2)$. The matrix $P^{(n)}$ is obtained from P in Eq. (3.19) by considering a piece-wise constant approximation for $\mathbf{q}(t)$, i.e., assuming that $\mathbf{q}(t) \simeq \mathbf{q}^{(n)}$ for $t \in (t_n-\delta/2, t_n+\delta/2]$, with $\mathbf{q}^{(n)} \triangleq \mathbf{q}(t_n)$. Specifically,

$$P^{(n)} = \begin{pmatrix} -j\lambda & q_1^{(n)}(t) & \dots & q_M^{(n)}(t) \\ -\sigma q_1^{(n)*}(t) & j\lambda & & \\ \vdots & & \ddots & \\ -\sigma q_M^{(n)*}(t) & & & j\lambda \end{pmatrix}. \tag{3.73}$$

The starting point, given by the boundary condition for $\boldsymbol{\phi}(t,\lambda)$ in $t = -T-\delta/2$, is

$$\boldsymbol{\phi}^{(1)} = \begin{pmatrix} 1 \\ \mathbf{0}_{M\times 1} \end{pmatrix} e^{j\lambda(T+\delta/2)}. \tag{3.74}$$

The scattering data are obtained from the end point solution as

$$a(\lambda) = \phi_1^{(N_t+1)} e^{+j\lambda(T+\delta/2)} \tag{3.75}$$

and

$$b_m(\lambda) = \phi_{m+1}^{(N_t+1)} e^{-j\lambda(T+\delta/2)}, \tag{3.76}$$

for $m = 1, \dots, M$.

The solution of the Cauchy problem (3.72) is obtained by using the transfer-matrix approach [7]. For each iteration (elementary step in t) we have $\boldsymbol{\phi}^{(n+1)} = U^{(n)}\boldsymbol{\phi}^{(n)}$, where $U^{(n)} = \exp(P^{(n)}\delta)$ is the transfer matrix. A closed form for the transfer matrix is derived in the next paragraphs. The desired multidimensional scattering data for the VNLSE (defining our NFT spectrum) can be obtained as

$$\begin{cases} a(\lambda) = \Sigma_1 e^{j\lambda(2T+\delta)}, \\ b_m(\lambda) = \Sigma_{m+1}, \end{cases} \quad \text{for } m = 1, \dots, M \tag{3.77}$$

where

$$\boldsymbol{\Sigma} = U^{(N_t+1)} \cdots U^{(1)} \begin{pmatrix} 1 \\ \mathbf{0}_{M\times 1} \end{pmatrix} \in \mathbb{C}^{M+1\times 1}. \tag{3.78}$$

Moreover, denoting by the prime the derivative with respect to λ, the coefficient $a'(\lambda)$ (which is used for the computation of the norming constants) is obtained as

$$a'(\lambda) = \phi_1^{'(N_t+1)} e^{j\lambda 2T} \tag{3.79}$$

where $\boldsymbol{\phi}'^{(N_t+1)}$ is computed from the recursion

$$\boldsymbol{\phi}'^{(n+1)} = U^{(n)}\boldsymbol{\phi}'^{(n)} + U'^{(n)}\boldsymbol{\phi}^{(n)}, \tag{3.80}$$

initialized by setting $\boldsymbol{\phi}'^{(1)} = (0, \dots, 0)^{\mathrm{T}}$.

Approximation to the transfer matrix From the definition of matrix exponential

$$U^{(n)} = \exp(P^{(n)}\delta) = \sum_{k=0}^{+\infty} \frac{1}{k!}(P^{(n)}\delta)^k \tag{3.81}$$

and neglecting the non-diagonal components of the squared matrix $P^{(n)2}$ ($P^{(n)2}$ is a diagonal matrix only for the scalar NLSE case with $M = 1$), by using the Taylor expansion for sinh and cosh functions, and doing some straightforward calculations, we obtain the following expression for the single-step transfer matrix

$$U^{(n)} \simeq U_a^{(n)} = \left(\begin{array}{c|ccc} c_0 - j\lambda s_0 & q_1^{(n)} s_0 & \dots & q_M^{(n)} s_0 \\ \hline -\sigma q_1^{(n)*} s_1 & c_1 + j\lambda s_1 & & \\ \vdots & & \ddots & \\ -\sigma q_M^{(n)*} s_M & & & c_M + j\lambda s_M \end{array}\right), \tag{3.82}$$

where $c_k = c_k^{(n)} = \cosh(\delta d_k)$ and $s_k = s_k^{(n)} = \sinh(\delta d_k)/d_k$ for $k = 0, 1, \dots, M$, with

$$d_0 = d_0^{(n)} = \sqrt{-\lambda^2 - \sigma \sum_{k=1}^{M} |q_k^{(n)}|^2}, \tag{3.83}$$

and

$$d_k = d_k^{(n)} = \sqrt{-\lambda^2 - \sigma |q_k^{(n)}|^2} \quad \text{for } k = 1, \dots, M. \tag{3.84}$$

Moreover,

$$U'^{(n)} \simeq U_a'^{(n)} = \left(\begin{array}{c|ccc} \Theta_0 & -q_1^{(n)}\lambda/d_0^2\ell_0 & \dots & -q_M^{(n)}\lambda/d_0^2\ell_0 \\ \hline \sigma q_1^{(n)*}\lambda/d_1^2\ell_1 & \Theta_1 & & \\ \vdots & & \ddots & \\ \sigma q_M^{(n)*}\lambda/d_M^2\ell_M & & & \Theta_M \end{array}\right), \tag{3.85}$$

where $\Gamma_0 = \Gamma_0^{(n)} = (\lambda\delta + j + j\lambda^2/d_0^2)$, $\Theta_0 = \Theta_0^{(n)} = j\lambda^2\delta/d_0^2 c_0 - \Gamma_0 s_0$, $\Gamma_m = \Gamma_m^{(n)} = (-\lambda\delta + j + j\lambda^2/d_m^2)$ and $\Theta_m = \Theta_m^{(n)} = -j\lambda^2\delta/d_m^2 c_m + \Gamma_m s_m$ for $m = 1, \dots, M$, and $\ell_m = (\delta c_m - s_m)$ for $m = 0, \dots, M$.

On the one hand, in the single-polarization case $M = 1$, the derivation is exact and was already given in [11, 58]. On the other hand, when $M \geq 2$, the transfer matrices $U^{(n)}$ Eq. (3.82) and $U'^{(n)}$ Eq. (3.85) are derived neglecting the non diagonal terms of the matrix $P^{(n)2}$ (and therefore, of all the $P^{(n)k}$ for $k \geq 2$), and, thus, are not exact. In particular, let $P^{(n)2} = D^2 + A^2$ where D^2

is a diagonal matrix and A^2 is a non diagonal matrix, then

$$\begin{aligned} U^{(n)} &= \mathbf{I} + P^{(n)}\delta + \sum_{k\geq 2} \frac{(P^{(n)2})^{k/2}\delta^k}{k!} \\ &= \underbrace{\mathbf{I} + P^{(n)}\delta + \sum_{k\geq 2} \frac{\delta^k}{k!}(D^2)^{k/2}}_{U_a^{(n)}} + \underbrace{\sum_{k\geq 2} \frac{\delta^k}{k!}\left[\sum_{\ell=0}^{k/2-1} \binom{k/2}{\ell} (D^2)^{\ell}(A^2)^{k/2-\ell}\right]}_{B^{(n)}} \\ &= U_a^{(n)} + B^{(n)} \end{aligned} \tag{3.86}$$

However, since δ is small, $B^{(n)} = o(\delta)$ and in the first order approximation

$$U(q_n) = U_a^{(n)} + B^{(n)} \simeq U_a(q_n). \tag{3.87}$$

The simulation results reported in the Chapters 4 and 5 have been obtained with this method. In particular concerning those in Chapter 5 for the ME, we verified that the use of a non-exact method does not hamper the results shown (see Figs. 5.6(a)-(b)).

Exact transfer matrix An exact closed form for the transfer matrix $U^{(n)}$ can be obtained via eigenvalue decomposition as described in Section A.2. The components of the transfer matrix are

$$U_{m,\ell}^{(n)} = \begin{cases} c - j\lambda s & m = \ell = 1 \\ q_{\ell-1}^{(n)} s & m = 1,\ \ell \geq 2 \\ -\sigma q_{m-1}^{(n)*} s & m \geq 2,\ \ell = 1 \\ r_{m-1,\ell-1}\left[c + j\lambda s - e^{j\lambda\delta}\right] & m = 2,\ \ell \geq 3 \text{ or } \ell = 2,\ m \geq 3 \\ r_{m-1,\ell-1}\left[c + j\lambda s\right] + e^{j\lambda\delta}(1 - r_{m-1,m-1}) & m = \ell = 2 \text{ or } m, \ell \geq 3 \end{cases} \tag{3.88}$$

where

$$\begin{aligned} c &= c^{(n)} = \cosh(d\delta), \\ s &= s^{(n)} = \sinh(d\delta)/d, \\ d &= d^{(n)} = \sqrt{-\lambda^2 - \sigma \textstyle\sum_{k=1}^{M} |q_k^{(n)}|^2} \\ r_{m\ell} &= r_{m\ell}^{(n)} = q_m^{(n)*} q_\ell^{(n)} / \textstyle\sum_{k=1}^{M} |q_k^{(n)}|^2 \end{aligned} \quad . \tag{3.89}$$

The inverse of the transfer matrix, which can be used for the computation of the solution of the Z-S problem starting from $+\infty$ rather than $-\infty$, is

$$U_{m,\ell}^{(n)-1} = \begin{cases} c + j\lambda s & m = \ell = 1 \\ -q_{\ell-1}^{(n)} s & m = 1,\ \ell \geq 2 \\ \sigma q_{m-1}^{(n)*} s & m \geq 2,\ \ell = 1 \\ r_{m-1,\ell-1}\left[c - j\lambda s - e^{-j\lambda\delta}\right] & m = 2,\ \ell \geq 3 \text{ or } \ell = 2,\ m \geq 3 \\ r_{m-1,\ell-1}\left[c - j\lambda s\right] + e^{-j\lambda\delta}(1 - r_{m-1,m-1}) & m = \ell = 2 \text{ or } m, \ell \geq 3 \end{cases} . \tag{3.90}$$

Moreover, the components of $U^{'(n)}$are

$$U^{'(n)}_{m,\ell} = \begin{cases} j\frac{\lambda^2}{d^2}\delta c - \Gamma^+ s & m = \ell = 1 \\ -q^{(n)}_{\ell-1}\frac{\lambda}{d^2}(\delta c - s) & m = 1,\ \ell \geq 2 \\ \sigma q^{(n)*}_{m-1}\frac{\lambda}{d^2}(\delta c - s) & m \geq 2,\ \ell = 1 \\ r_{m-1,\ell-1}\left(-j\frac{\lambda^2}{d^2}\delta c + \Gamma^- s - j\delta e^{j\lambda\delta}\right) & m = 2,\ \ell \geq 3 \text{ or } \ell = 2,\ m \geq 3 \\ r_{m-1,\ell-1}\left(-j\frac{\lambda^2}{d^2}\delta c + \Gamma^- s\right) + j\delta e^{j\lambda\delta}(1 - r_{m-1,m-1}) & m = \ell = 2 \text{ or } m, \ell \geq 3 \end{cases} \tag{3.91}$$

where $\Gamma^{\pm} = j + j\lambda^2/d^2 \pm \lambda\delta$.

Importantly, for the ME, i.e., $M = 2$, the transfer matrix reduces to

$$U^{(n)} = \begin{pmatrix} c - j\lambda s & q^{(n)}_1 s & q^{(n)}_2 s \\ -\sigma q^{(n)*}_1 s & r_{1,1}\left[c + j\lambda s - e^{j\lambda\delta}\right] + e^{j\lambda\delta} & r_{1,2}\left[c + j\lambda s - e^{j\lambda\delta}\right] \\ -\sigma q^{(n)*}_2 s & r_{2,1}\left[c + j\lambda s - e^{j\lambda\delta}\right] & r_{2,2}\left[c + j\lambda s - e^{j\lambda\delta}\right] + e^{j\lambda\delta} \end{pmatrix}. \tag{3.92}$$

Comparison between approximated and exact methods For $M = 1$ (single-polarization NLSE) both the exact (3.88) and approximated (3.82) transfer matrices become

$$U^{(n)} = \begin{pmatrix} c - j\lambda s & q^{(n)}_1 s \\ -\sigma q^{(n)*}_1 s & c + j\lambda s \end{pmatrix}, \tag{3.93}$$

as in [11, 58]. Moreover, notice that considering $q_2(t) = 0$ in Eq. (3.92), one obtains

$$U^{(n)} = \begin{pmatrix} c - j\lambda s & q^{(n)}_1 s & 0 \\ -\sigma q^{(n)*}_1 s & c + j\lambda s & 0 \\ 0 & 0 & e^{j\lambda\delta} \end{pmatrix}, \tag{3.94}$$

which produces the same results as for $M = 1$, since the third component is initialized with 0.

On the one hand, for the comparison of the two methods in terms of performance, we refer to the end Section 5.1 and Fig. 5.6 for the special case with $M = 2$, while the case with $M \geq 3$ will be considered in a future work. On the other hand, as far as it concerns the computational complexity of the two algorithms, it is worth noting that multiplying the transfer matrix for a vector $\mathbf{y}$ as for Eq. (3.78), requires (excluding the operations to compute the matrices components)

$$3M + 1 \tag{3.95}$$

complex products with the approximated method and

$$(M + 1)^2 \tag{3.96}$$

complex products for the exact method, making the latter much more involved. Of course, for a proper comparison between the two methods, one should consider the computational cost together with the accuracy, which has not been tested for $M \geq 3$. We postpone this analysis to a future work.

3.4.2 Numerical backward nonlinear Fourier transform via Gelfand-Levitan-Marchenko equation

This subsection derives the NCG method to compute the BNFT for the VNLSE [40], generalizing the concepts used for the NLSE in [41, 65]. The method presented applies to the VGLME of arbitrary dimension M and can be used in presence of both the discrete and continuous spectrum.

The BNFT is computed by solving the VGLME, Eq. (3.29), of dimension $1 \times M$. This equation can be rewritten as the Marchenko system of order $M+1$:

$$\begin{cases} B_1(t,\alpha) - \sigma \int_0^{+\infty} \mathbf{B}_2(t,\beta)\mathbf{F}(\alpha+\beta+2t)\, d\beta = 0 \\ \mathbf{B}_2(t,\alpha) - \sigma\mathbf{F}^\dagger(\alpha+2t) + \int_0^{+\infty} B_1(t,\beta)\mathbf{F}^\dagger(\alpha+\beta+2t)\, d\beta = 0 \end{cases} \tag{3.97}$$

where $B_1(x,\alpha)$ is a scalar unknown function and $\mathbf{B}_2(x,\alpha)$ is a vector of $1 \times M$ unknown functions. The first equation is a scalar equation, while the second one contains M equations. The time domain signal (of M components) is obtained as $\mathbf{q}(t) = -2\mathbf{B}_2(t,0)$.

Assuming that $|\mathbf{q}(t)| = 0$ for $|t| > T$, consider the uniform grid over the interval $[-T,T]$, with $t_k = -T + (k-1)\delta$ for $k = 1,\dots,N_t+1$ and time step $\delta = 2T/N_t$. For each $t_k \in [-T,T]$, the solution $\mathbf{q}(t_k) = -2\mathbf{B}_2(t_k,0)$ depends on the values of $\mathbf{F}(\alpha+\beta+2t_k)$, $\mathbf{B}_2(t_k,\beta)$, and $B_1(t_k,\beta)$ for $\alpha,\beta \geq 0$, as clear from Eq. (3.97). Therefore, we shall consider a uniform grid for α and β in the interval $[0,T_B)$ (given T_B such that $|\mathbf{F}(y)| = 0$ for $y \geq 2t_k + T_B$, and $|B_1(t_k,\alpha)| = 0$ and $|\mathbf{B}_2(t_k,\alpha)| = 0$ for $\alpha \geq T_B$) with discretization step 2δ: $\alpha_\ell = (l-1)2\delta$ and $\beta_\ell = (\ell-1)2\delta$ for $\ell = 1,\dots,L$, with $L = \lceil T_B/(2\delta)\rceil$. Using the Nyström method, the Marchenko system in Eq. (3.97) can be reduced to a linear system, discretizing the integrals through the composite Simpson's quadrature rule with weights $\mathbf{d}$ (see Section A.5).

The linear system equivalent to (3.97) is

$$\underbrace{\left(\begin{array}{c|ccc} \mathbf{I}_L & -\sigma H_1 D & \dots & -\sigma H_M D \\ \hline H_1^* D & & & \\ \vdots & & \mathbf{I}_{LM} & \\ H_M^* D & & & \end{array}\right)}_{L(M+1)\times L(M+1)} \underbrace{\begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_{2,1} \\ \vdots \\ \mathbf{b}_{2,M} \end{pmatrix}}_{L(M+1)\times 1} = \underbrace{\begin{pmatrix} \mathbf{0}_{L\times 1} \\ \sigma\mathbf{f}_1^* \\ \vdots \\ \sigma\mathbf{f}_M^* \end{pmatrix}}_{L(M+1)\times 1}, \tag{3.98}$$

where D is the $L \times L$ diagonal matrix that defines the quadrature rule according to $D_{\ell,\ell} = d_\ell$ for $\ell = 1,\dots,L$; $\mathbf{b}_1$ and $\mathbf{b}_{2,m}$ are the $L \times 1$ vectors containing, respectively, the values of $B_1(t_k,\alpha)$ and $\mathbf{B}_{2,m}(t_k,\alpha)$, i.e., $b_{1\ell} = B_1(t_k,\alpha_\ell)$ and $b_{2,m,\ell} = B_{2,m}(t_k,\alpha_\ell)$ for $\ell = 1,\dots,L$; H_m is the $L \times L$ matrix containing the values of $F_m(\alpha+\beta+2t_k)$ with $H_{m,\ell,g} = F_m(\alpha_\ell+\alpha_g+2t_k)$ for $\ell,g = 1,\dots,L$; and $\mathbf{f}_m$ is the vector $L \times 1$ containing the values of $F_m(\alpha+2t_k)$ with $f_{m,\ell} = F_m(\alpha_\ell+2t_k)$ for $\ell = 1,\dots,L$. Importantly, $\mathbf{f}_m$ is the first row of the matrix H_m, and H_m is the triangular upper left Hankel matrix generated by the vector $\mathbf{f}_m$ (see Eq. (2.13)).

Eq. (3.98) can be equivalently written in a compact form as

$$\underbrace{\begin{pmatrix} \mathbf{I}_L & -\sigma\mathbf{H}\mathbf{D}_{LM} \\ \mathbf{H}^\dagger D & \mathbf{I}_{LM} \end{pmatrix}}_{L(M+1)\times L(M+1)} \underbrace{\begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix}}_{L(M+1)\times 1} = \underbrace{\begin{pmatrix} \mathbf{0}_{L\times 1} \\ \sigma\mathbf{f}^* \end{pmatrix}}_{L(M+1)\times 1}, \tag{3.99}$$

where $\mathbf{H} = (H_1, \ldots, H_m)$ is an $L \times LM$ matrix, $\mathbf{f} = (\mathbf{f}_1^{\mathrm{T}}, \ldots, \mathbf{f}_m^{\mathrm{T}})^{\mathrm{T}}$ is a $LM \times 1$ vector, $\mathbf{b}_2 = (\mathbf{b}_1^{\mathrm{T}}, \ldots, \mathbf{b}_M^{\mathrm{T}})^{\mathrm{T}}$ is a $LM \times 1$ vector, and $\mathbf{D}_{LM}$ is the $LM \times LM$ diagonal matrix with diagonal $(\mathbf{d}, \ldots, \mathbf{d})$. Substituting the first row into the second and multiplying by $\mathbf{D}_{LM}$, we obtain the system of LM equations:

$$(\mathbf{D}_{LM}\mathbf{H}^\dagger D\mathbf{H}\mathbf{D}_{LM} + \sigma\mathbf{D}_{LM})\mathbf{b}_2 = \mathbf{D}_{LM}\mathbf{f}^*, \tag{3.100}$$

from which the solution at the time instant t_k is obtained as $q_m(t_k) = -2b_{2,(m-1)L}$.

The last equation is the analog of that derived in [65] for the scalar GLME (i.e., when $M = 1$), where it is numerically solved using the conjugate gradient method by taking advantage of the fact that the system's matrix is symmetric and positive-defined, and of the Hankel shape of the matrices involved. Unfortunately, while the conjugate gradient method can also be used in our case, the matrix $\mathbf{H}$ in Eq. (3.100) is not Hankel, and therefore, the matrix multiplication may be a computationally demanding task for our problem. However, Eq. (3.100) is equivalent to the following system of equations:

$$A_{m,1}\mathbf{b}_{2,1} + \cdots + A_{m,M}\mathbf{b}_{2,M} + \sigma D\mathbf{b}_{2,m} = D\mathbf{f}_m^*, \tag{3.101}$$

for $m = 1, \ldots, M$, where $A_{m,n} = DH_m^\dagger DH_n D$, and H_m-s are the Hankel matrices. Consequently, the system (3.100) can now be solved with the conjugate gradient method through Eq. (3.101), starting with an initial guess for $\mathbf{b}_2$ (e.g., the null vector) and iteratively updating the solution, and performing the products involved with help of the fast Fourier transform (FFT) as explained in Section A.4.

The method explained above should be independently applied to find the solution in any time instant t_k of interest. However, if the solution has to be found in the whole interval $[-T, T]$, several iterations can be saved starting from $t_{N_t+1} = T$, and later for t_k considering as a starting point for $\mathbf{b}_2$ the vector found in the previous step, at the adjacent time instant t_{k+1}.

In this work, we considered the nonlinear spectrum *from the right* defined as $\boldsymbol{\rho}(\lambda) = \mathbf{b}(\lambda)/a(\lambda)$ [7] (i.e. the right reflection coefficient), and the corresponding VGLME given by Eq. (3.29). However, one can also consider the nonlinear spectrum *from the left* $\boldsymbol{\rho}_l(\lambda) = \mathbf{b}(\lambda)^*/a(\lambda)$ [7] and its corresponding VGLME, which is different from Eq. (3.29) but can be obtained from it [65]. The authors of [65], considering the scalar NLSE case only, explain that while from a theoretical point of view it is equivalent to use one nonlinear spectrum (left or right) instead of the other, from the numerical point of view, the accuracy of the numerical method can be significantly improved by considering the standard GLME from the right to find the time domain signal in time instants $t_k \geq 0$,

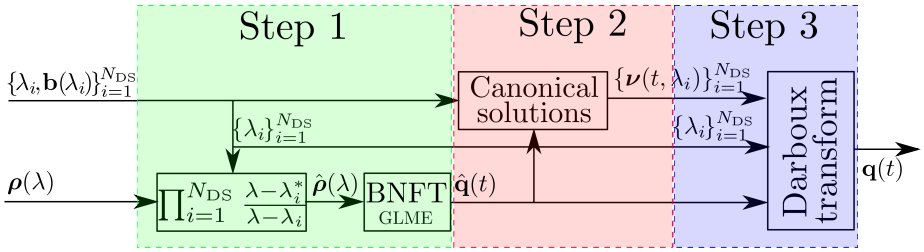


Figure 3.3: The BNFT for joint spectrum through GLME and DT.

and the GLME from the left for $t_k < 0$. We expect that the same should hold for the VGLME. However, in our work we used only the standard VGLME from the right (3.29).

The method presented here can, in principle, be applied to any nonlinear spectrum. However, the method might not converge when the energy of the signal is too high: in this case, the method starts to diverge for increasing values of t starting from $-\infty$, showing some characteristic peaks. This situation does not occur in practical scenarios when only the continuous spectrum is considered. Instead, it can significantly affects the method when discrete components are also present, as in this case the energy of the signal is higher. Several solutions can be considered to overcome this issue. Firstly, we expect that when considering the NCG method combining both the nonlinear spectrum from the left and from the right, this problem would be postponed to higher energy levels, in practice solving the problem (also when considering discrete eigenvalue modulation). Indeed, in this case, the values of the optical signal for $t < 0$, where the problem occurs, are obtained from the nonlinear spectrum from the left, which, on the other hand, causes problems for decreasing values of t starting from $+\infty$. Secondly, other numerical methods might be more robust against this effect, e.g., the Frumin method [66] for the scalar NLSE, which also solves the GLME, is affected by the same issue but for much higher energy values. However, an extension to the multidimensional case has not been derived yet. Finally, since this problem affects systems of practical interest for our purposes mostly when the discrete part is non zero, it can be avoided by combining the NCG with the DT, as described in the following section.

3.4.3 Numerical backward nonlinear Fourier transform via Gelfand-Levitan-Marchenko equation and Darboux transform

The DT [53, 67] is a method to add discrete eigenvalues to a given time domain signal. The method was firstly proposed for optical communication purposes in [11] to obtain the time domain signal corresponding to multi solitons (no continuous spectrum) for the NLSE, and later extended for the same purpose to the ME in [68, 70, 71]. However, the DT can also be applied to add dis-

crete components to a given time domain signal having non-zero continuous nonlinear spectrum. Consequently, one can perform the BNFT to obtain the time domain signal $\mathbf{q}(t)$ given a continuous spectrum $\boldsymbol{\rho}(\lambda)$ and a discrete spectrum $\{\lambda_i, \boldsymbol{C}_i\}_{i=1}^{N_{\mathrm{DS}}}$ with three steps, as described and proved for the scalar NLSE in [14]. The procedure was extended to the ME (i.e., the VNLSE with $M = 2$) in [26], and is described in this Subsection. Importantly, this procedure has not been proved for dual-polarization, but merely extended from single-polarization and later verified a posteriori with numerical simulations, as will be shown in Section 6.1.

Given the joint dual-polarization spectrum $\boldsymbol{\rho}(\lambda)$ and$\{\lambda_i, \mathbf{b}(\lambda_i)\}_{i=1}^{N_{\mathrm{DS}}}$, the corresponding time domain signal $\mathbf{q}(t)$ can be obtained with the following three-step algorithm, sketched in Fig.3.3[3].

1. Compute the time domain signal $\hat{\mathbf{q}}(t)$ corresponding to the modified continuous spectrum
$$\hat{\boldsymbol{\rho}}(\lambda) = \boldsymbol{\rho}(\lambda) \prod_{i=1}^{N_{\mathrm{DS}}} \frac{\lambda - \lambda_i}{\lambda - \lambda_i^*}. \tag{3.102}$$
This can be achieved with the NCG method (See Subsection 3.4.2).

2. For each eigenvalue λ_i for $i = 1, \ldots, N_{\mathrm{DS}}$, obtain the solution $\boldsymbol{\nu}(t, \lambda_i)$ of the eigenvalue problem $L(\hat{\mathbf{q}})\boldsymbol{\nu} = \lambda_i \boldsymbol{\nu}$ associated with $\hat{\mathbf{q}}$ with boundary conditions
$$\begin{cases} \nu_1(T, \lambda_i) = 1 \\ \nu_2(-T, \lambda_i) = -b_1(\lambda_i) \\ \nu_3(-T, \lambda_i) = -b_2(\lambda_i) \end{cases} \tag{3.103}$$
where $\hat{\mathbf{q}}(t) = 0$ for $t \notin [-T, T]$. The solution $\boldsymbol{\nu}$ can be obtained as
$$\boldsymbol{\nu}(t, \lambda_i) = \frac{1}{\phi_1(T, \lambda_i)} \boldsymbol{\phi}(t, \lambda_i) - \boldsymbol{\psi}(t, \lambda_i) \boldsymbol{\psi}^{(2)}(-T, \lambda_i)^{-1} \mathbf{b}(\lambda_i), \tag{3.104}$$
where $\boldsymbol{\phi}(t, \lambda_i)$ and $\boldsymbol{\psi}(t, \lambda_i)$ are the conventional solutions of the eigenvalue problem $L(\hat{\mathbf{q}})\mathbf{v} = \lambda \mathbf{v}$ defined by the boundary conditions Eq.s (3.20)-(3.21), i.e.,
$$\boldsymbol{\phi}(-T, \lambda_i) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} e^{j\lambda_i T} \text{ and } \boldsymbol{\psi}(T, \lambda_i) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} e^{j\lambda_i L} \tag{3.105}$$
and $\boldsymbol{\psi}^{(2)}(t, \lambda_i)$ is the 2×2 matrix made of the second and the third rows of $\boldsymbol{\psi}(t, \lambda_i)$, i.e.,
$$\boldsymbol{\psi}^{(2)}(t, \lambda_i) = \begin{pmatrix} \psi_{21} & \psi_{22} \\ \psi_{31} & \psi_{32} \end{pmatrix}. \tag{3.106}$$
The conventional solutions $\boldsymbol{\phi}(t, \lambda_i)$ and $\boldsymbol{\psi}(t, \lambda_i)$ can be found with standard methods for the FNFT, as described in Section 3.4.1.

[3]Here, the method is derived starting from $\mathbf{b}(\lambda_i)$, rather than considering the norming constant$\mathbf{C}_i$, as $\mathbf{b}$-modulation has been recently shown to perform better for the discrete spectrum [72].

3. Execute the DT for the ME [68] with input parameters $\hat{\mathbf{q}}(t)$, $\{\lambda_i\}_{i=1}^{N_{\mathrm{DS}}}$, and $\{\boldsymbol{\nu}(t,\lambda_i)\}_{i=1}^{N_{\mathrm{DS}}}$ being the generic auxiliary solutions to iteratively add the discrete eigenvalues $\{\lambda_i, \mathbf{b}(\lambda_i)\}_{i=1}^{N_{\mathrm{DS}}}$ to $\hat{\mathbf{q}}(t)$. The solution $\mathbf{q}(\lambda)$ obtained from the DT has continuous spectrum $\boldsymbol{\rho}(\lambda)$ and discrete spectrum $\{\lambda_i, \mathbf{b}(\lambda_i)\}_{i=1}^{N_{\mathrm{DS}}}$, as verified with numerical simulations in Chapter 6.

3.5 The nonlinear Fourier transform in optical fiber communication

This Chapter showed that the NFT is a powerful tool to solve different flavors of the VNLSE, which are of practical interest for the opt-com community as master models for the propagation of the optical signal in optical fibers, as shown in Chapter 2. Consequently, the NFT can be an effective tool for fiber optic communication: (i) it can be used for DBP, in practice replacing the SSFM with a single tap operation; (ii) it can be used to encode the information on the nonlinear spectrum; or (iii) it can be used for hybrid methods [7]. However, in the past years, approaches borrowed from linear communication have been effectively used for fiber optic communication, in practice replacing NFT-based methods. Only recently (except for the early work of Hasegawa and Nyu [22]), researchers started to consider transmission schemes based on the NFT.

On the one hand, the use of the NFT for DBP [7, 56] is, in principle, more effective than the SSFM as it allows to remove deterministic propagation effects with a single tap operation, i.e., by multiplying the received nonlinear spectrum for the inverse of the channel filter in Prop. 6, rather than performing several FFTs as for the SSFM. However, this would require to perform one FNFT and one BNFT, and, unfortunately, fast numerical methods for the NFT (i.e., comparable to the FFT) are still subject of research activity (see Section 3.4). On the other hand, in the past years, research has been mostly focused on transmission paradigms that encode the information on the nonlinear spectrum—referred to as NFDM. Moreover, due to the pioneering nature of the topic, researchers mostly focused on the NFT applied to the scalar NLSE with vanishing boundary conditions; with only a few works about periodic boundary conditions [17,52], and, only recently, several works about dual-polarization schemes [26, 40, 49, 68, 71, 73, and others].

This thesis is focused on NFDM schemes, employing the NFT with vanishing boundary conditions for information encoding and decoding. We refer to Chapter 4 for single-polarization NFDM and to Chapters 5-6 for dual-polarization NFDM.

3.6 Conclusion and outlook

This Chapter dealt with the NFT for optical communication. We reviewed part of the theory and concepts related to the NFT that are important for their application to optical fiber communication, i.e., for the multidimensional

localized VNLSE. Furthermore, we reported some numerical methods for the computation of the FNFT and the BNFT, aiming at providing some general (despite not optimal) methods.

The optimization of numerical methods for the NFT is a topic of active research, essential for the application of the NFT to optical communication. Furthermore, the development of fast numerical methods for the NFT for the NLSE with periodic boundary conditions [17,52] is required for the development of transmission paradigms not employing burst transmission.

Chapter 4

Single-polarization nonlinear frequency-division multiplexing systems

This chapter investigates, through simulations, some transmission schemes for optical fiber communication based on NFDM—a novel transmission paradigm which uses the NFT to master fiber nonlinearity and potentially overcome the limitations of conventional systems. Due to the pioneering nature of the topic and despite the growing interest in the opt-com community, the optimization and understanding of the topic is not trivial and requires a huge amount of work on different aspects and research fields.

With this in mind, this chapter aims to investigate NFDM in a simple scenario, which, however, requires a non trivial analysis. In particular, this Chapter focuses on single-polarization NIS—a popular NFDM scheme which considers localized signals and do not include discrete spectrum modulations—considering the lossless fiber channel (i.e., assuming ideal Raman amplification) and the single-user scenario. Different mappings are briefly considered in the end of this Chapter, while the utilization of both polarizations to double the information rate will be considered in the next Chapter. Localized signals, i.e., with vanishing boundary conditions, are considered for the reasons already explained in Section 3.5. Furthermore, this Chapter considers only the modulation of the continuous spectrum which, we believe, is essential to achieve high spectral efficiency; eventually, it can be combined with discrete spectrum modulation as in [14, 18] and Chapter 6. As far as it concerns the lossless model, required for the NFT-integrability, it represents a good approximation when the LPA model is considered without significantly degrading performance [30]. The effect of loss on a two-soliton system has been investigated in [74]. Finally, we consider the single-user scenario because, while NFT-based schemes would grant the best improvements with respect to conventional systems in the multi-user scenario, we believe that the optimization of the single-user scenario is essential to achieve

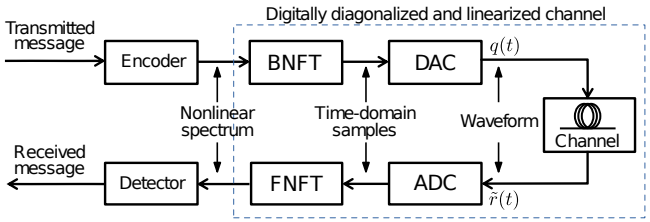


Figure 4.1: A typical NFDM scheme.

high spectral efficiency and relevant results.

This Chapter is organized as follows. Section 4.1 gives a brief introduction about NFDM schemes, while Section 4.2 investigates the performance of a NIS scheme, with particular attention to numerical issues, highlights some important limitations of NIS, and compares its performance with currently deployed techniques (purely linear techniques using EDC and nonlinearity compensation techniques using DBP). Section 4.3 and Section 4.4 describes two techniques aimed at reducing some of the issues related to NIS. Section 4.5 proposes three different detection strategies for NIS which provide significant performance improvements. Finally, Section 4.6 briefly considers other type of mappings (different from NIS) recently proposed, and Section 4.7 draws the conclusions. This Chapter reports some of the results presented in [54, 75–79].

4.1 Nonlinear frequency-division multiplexing schemes

NFDM is the umbrella term for optical fiber transmission schemes which encode (and decode) information on (from) the nonlinear spectrum of the optical signal, employing the integrability of the NLSE—which governs the propagation of a single-polarization signal in a SMF—with the NFT method. A typical NFDM scheme is sketched in Fig. 4.1. The advantage of NFDM with respect to other schemes stems from the trivial propagation of the nonlinear spectrum along an optical fiber (see Prop. 6). As an example, Fig. 4.2 shows the modulus of the optical signal, the linear spectrum, and the nonlinear spectrum at different length along the fiber: the time domain signals suffers from linear and nonlinear distortions (dispersion is particularly evident in this case), the linear spectrum suffers from nonlinear distortions, while the modulus of the nonlinear spectrum is invariant. NFDM is a sort of nonlinear analog of conventional OFDM, using the NFT rather than the FT. In this way, information is encoded on the nonlinear spectrum, i.e., on non-interacting degrees of freedom [12].

Different kinds of NFDM schemes have been proposed. *Eigenvalue communication* proposed by Hasegawa and Nyu [22] in 1993 was, in fact, the first NFDM paradigm. Then, considering vanishing boundary conditions NFT and single-polarization systems, several other NFDM schemes have been proposed for mod-

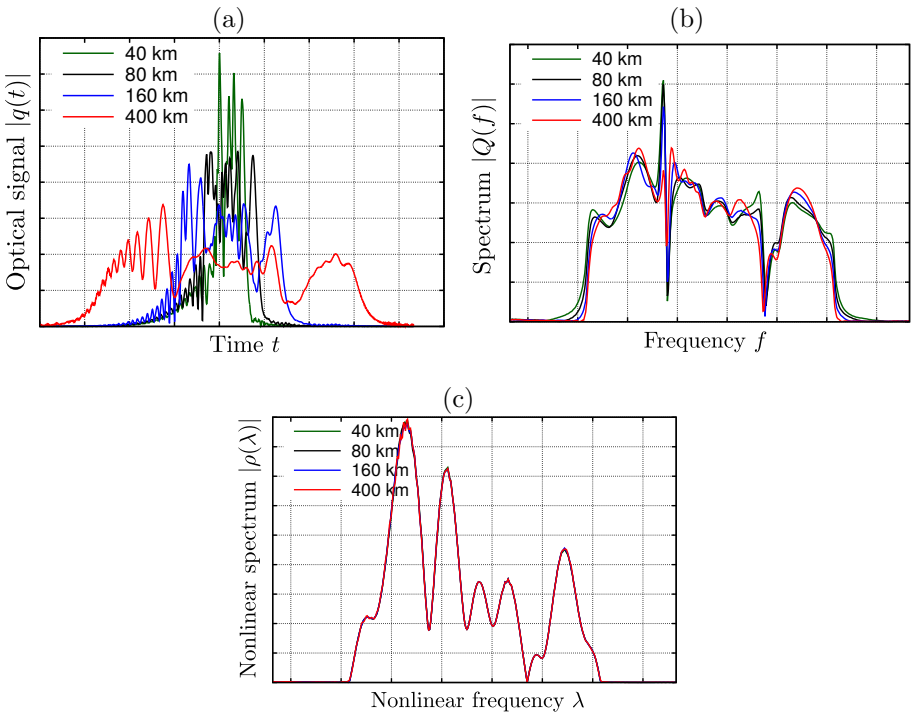


Figure 4.2: Evolution along the NLSE channel (noise-free and lossless fiber channel) of the modulus of the (a) optical signal; (b) linear spectrum; and (c) nonlinear spectrum.

ulating the continuous [12, 15, 16, 19, 30, 50, 64, 75], the discrete [11, 67, 80–83], or both parts of the nonlinear spectrum [14, 18, 84]. Many experimental demonstrations have been carried out [13, 14, 16, 18, 80, 82–84].

As far as it concerns the modulation of the continuous spectrum, different mappings have been proposed. The NIS technique [16, 19, 30, 50] maps the FT of a QAM signal carrying information (more in general, a linearly modulated signal) on the continuous nonlinear spectrum of the optical signal, such that at low power the optical signal tends to the QAM signal (more on this in the next section). Similarly, one can directly map the QAM signal, rather than its FT, on the continuous nonlinear spectrum. Furthermore, different mappings have been considered: *exponential* mapping and *b*-modulation. This Chapter is mainly focused on NIS, however, Subsection 4.6 briefly considers also the exponential mapping and the *b*-modulation.

4.2 Nonlinear inverse synthesis

The NIS technique [16, 19, 30, 50] is a popular NFDM scheme which maps the FT of a QAM signal on the nonlinear continuous spectrum of the optical signal; the modulation of the discrete part of the nonlinear spectrum is not considered in NIS, but can be added [14, 18]. The main advantages of NIS are that it can be combined with conventional coding and modulation to achieve high spectral efficiencies, it approaches channel capacity in the linear regime, and it can eventually be combined with discrete spectrum modulation.

The TX encodes the information on a QAM signal

$$s(t) = \sum_{k=1}^{N_b} x_k g(t - (k-1)T_s) \tag{4.1}$$

with pulse shape $g(t)$, symbol time T_s, and information symbols $x_1, \ldots, x_{N_b}$ chosen from a proper alphabet. After normalization (see Section 3.2.1), the FT $S(f)$ of $s(t)$ is mapped onto the continuous part of the input nonlinear spectrum $\rho(\lambda)$ according to

$$\rho(\lambda) = -S\left(-\lambda/\pi\right). \tag{4.2}$$

The normalized frequencies f are mapped on the nonlinear frequencies $\lambda = -\pi f$. Given $\rho(\lambda)$, the BNFT block generates the time-domain samples of the optical signal, which is then obtained with a digital-to-analog converter (DAC) and launched into the channel. At the RX, the output noisy optical signal $\tilde{r}(t)$ is sampled by the analog-to-digital converter (ADC) and sent to the FNFT block, which computes the corresponding output nonlinear spectrum $\tilde{\rho}(L, \lambda)$ (corrupted by amplifier noise during propagation). Finally, the detector multiplies $\tilde{\rho}(L, \lambda)$ by $e^{j4\lambda^2\mathcal{L}}$ to remove propagation effects and obtains a noisy version $\tilde{\rho}(\lambda)$ of the nonlinear spectrum (4.2) in which the information was encoded. Finally, the detector performs an inverse FT to recover a noisy version of $s(t)$, followed by matched filtering and symbol-time sampling to obtain a noisy replica

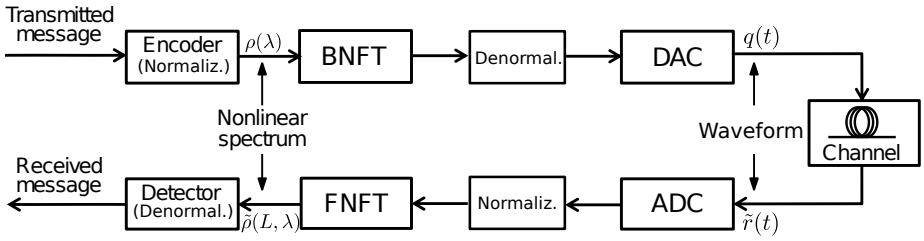


Figure 4.3: NIS scheme with normalization and denormalization procedures.

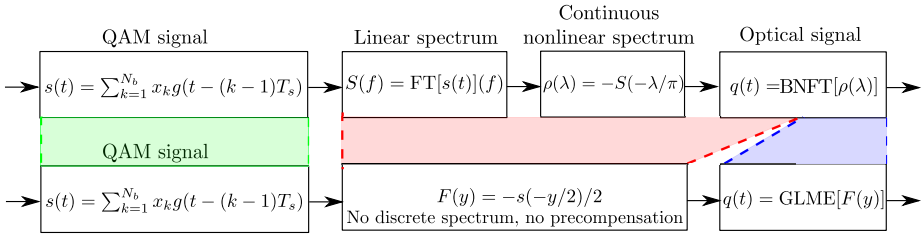


Figure 4.4: NIS mapping: full (upper part) and reduced (lower part).

of the transmitted symbols; and makes decisions based on a minimum Euclidean distance criterion. For more details about this detection strategy, refer to Section 4.5. Note that here and in the following description, for the sake of simplicity, normalization and denormalization operations are omitted, but they are required for both FNFT and BNFT. For the ease of computation, we can perform the digital operations with normalized variables, as depicted in Fig. 4.3.

Details of NIS mapping are shown in Fig. 4.4 in the upper part. Furthermore, when the discrete spectrum is empty and precompensation is not deployed (see Section 4.4) as in this case, NIS mapping can be realized using the GLME and directly mapping the QAM signal on $F(y)$ with

$$F(y) = -s(-y/2)/2 \tag{4.3}$$

as follows from (3.65), (4.1) and (4.2). The reduced NIS mapping is shown in the lower part in Fig. 4.4. Furthermore, when Eq. (4.3) holds, the optical signal tends to the QAM initial signal at low power as[1]

$$q(t) \to -\sigma s^*(-t). \tag{4.5}$$

Importantly, NIS transmission is organized in bursts of N_b information symbols, separated by a guard time (N_z spaces, duration $N_z T_s$) to fulfill (at least

[1]It follows, for $M \geq 1$, from the VGLME Eq. (3.29) and Eq. (4.3) in the asymptotics when $||\mathbf{q}||_1$ is small:

$$\mathbf{q}(t) = -2\mathbf{K}(t,t) \to -2\sigma\mathbf{F}^\dagger(2t) = -\sigma\mathbf{s}^\dagger(-t). \tag{4.4}$$

within the limit of channel memory) the vanishing boundary conditions of the underling NFT theory and avoid burst interaction during propagation. The guard time: *i)* plays a role similar to the cyclic prefix in OFDM; *ii)* should at least equal the maximum time broadening induced by fiber dispersion; and *iii)* causes a reduction of the overall spectral efficiency by the factor η—*rate efficiency* (2.17). The total duration of each burst is $T_s(N_b + N_z)$.

As it will be shown in this section, burst transmission causes some important drawbacks that, if not clearly understood and overcome, may severely limit performance and spectral efficiency. In fact, the simple solution (used in conventional OFDM) of increasing N_b to limit the loss of spectral efficiency is not feasible, because the performance decreases with N_b due to a sort of signal–noise interaction taking place at the receiver, as it will be shown in Section 4.2.1.2.

4.2.1 System performance: simulation setup and results

The performance of NIS has been investigated through simulations with the following setup. The QAM signal is a quadrature phase-shift keying (QPSK) signal with PSD raised-cosine shaped with roll-off factor $\beta = 0.2$ (a typical choice in conventional systems), while the symbol rate is $R_s = 1/T_s = 50\,\text{GBd}$. The physical channel is a SMF of length $L = 2000\,\text{km}$, attenuation $\alpha = 0.2\,\text{dB/km}$, GVD parameter $\beta_2 = -20.39\,\text{ps}^2/\text{km}$, and nonlinear coefficient $\gamma = 1.22\,\text{W}^{-1}\text{km}^{-1}$, along which ideal distributed amplification with spontaneous emission factor $\eta_{\text{sp}} = 4$ is considered. The bandwidth of both the DAC and the ADC is 100 GHz. The BNFT is computed with the NCG method, while the Boffetta–Osborne method is employed for the FNFT, both described in Section 3.4 with $M = 1$.

The performance of NIS is investigated measuring the Q-factor as a function of the launch power P_s in Eq. (2.19) for both the noise-free and the noisy scenarios. In the first case, numerical errors due to the use of a "discretized NFT" (more on this later) are the only source of noise and, as will be shown in the following, the performance decreases at higher powers due to the increasing impact of numerical inaccuracies. In the second case, at lower powers performance worsens (with respect to the noise-free scenario) due to the ASE noise which reduces the SNR. As will be shown in the following, when increasing the launch power (and, thus, the SNR), the performance increases up to an optimal point and then, it decreases again. Possible reasons of this decay are both numerical inaccuracies, similarly to the noise-free case, and signal–noise interactions, including signal and noise interacting in a nonlinear manner during propagation, signal–noise interaction at the FNFT, and noise affecting the integrability of the NLSE. The impact of these factors are investigated in the next Subsections.

4.2.1.1 Impact of discretization and boundary conditions in NIS

While the NFT is an exact theory, the algorithms and the techniques considered in this thesis (in general, for NFT-based communication), use a sort of "discretized NFT", obtained by discretizing the functions involved and assuming finite durations. As a consequence, we carried out extensive investigations

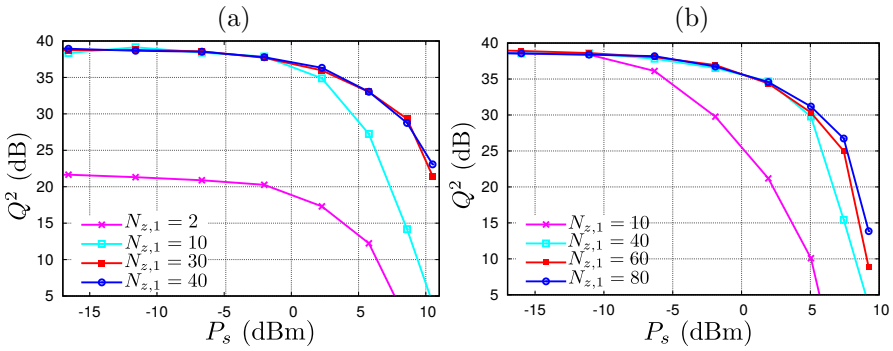


Figure 4.5: Performance in B2B configuration for $N_F = N_B = 4$ for different guard symbols $N_{z,1}$ and for (a) $N_b = 8$, and (b) $N_b = 16$. QPSK symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 800$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

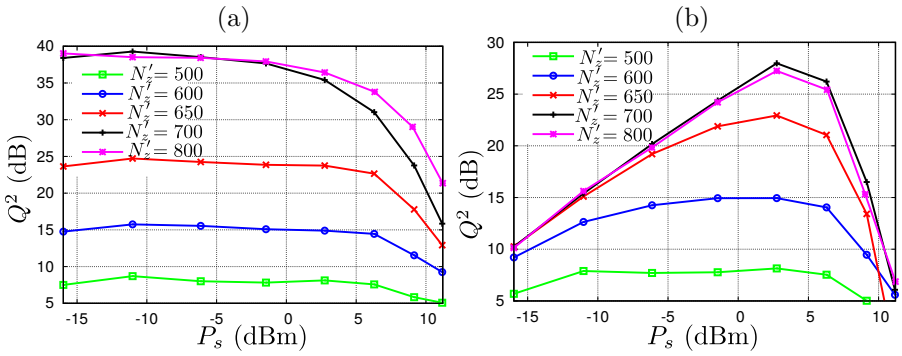


Figure 4.6: Performance for $N_b = 8$ and $N_F = N_B = 4$ for different N'_z considered for the FNFT processing at the RX in (a) noise-free scenario, and (b) noisy scenario. QPSK symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 8000$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

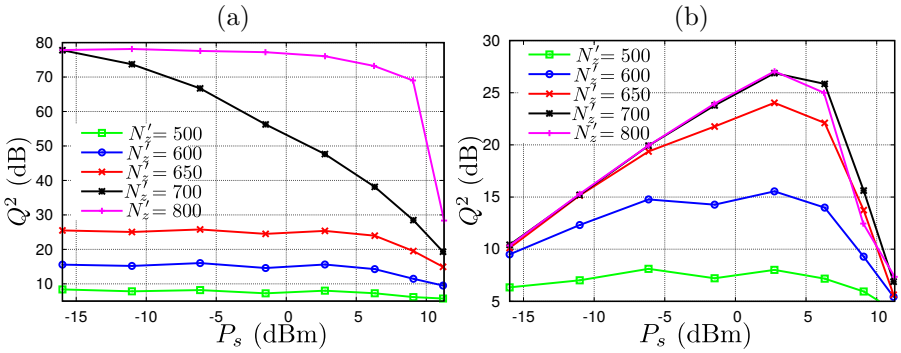


Figure 4.7: Performance for $N_b = 8$ and $N_F = N_B = 40$ for different N'_z considered for the FNFT processing at the RX in (a) noise-free scenario, and (b) noisy scenario. QPSK symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 800$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

to understand the impact of numerical discretization, i.e., oversampling factors for the NFT operations, and the fulfillment of the boundary conditions [78]. As far as it concerns the latter, the overall number of guard symbols is taken to be $N_z = 800$ according to the formula for the time broadening caused by GVD for a linear channel

$$\Delta T = 2\pi\beta_2 \Delta f L, \tag{4.6}$$

with Δf being the total bandwidth of the bandpass signal. However, computing the BNFT at the TX on the whole time window might be useless (as the signal is not yet broadened in time) while increasing the computational complexity; therefore, one can consider $N_{z,1} \leq N_z$ symbol time for the boundary condition of the BNFT, and then add additional $N_{z,2}$ symbol time to obtain $N_z = N_{z,1} + N_{z,2}$. Finally, the RX processes a portion of the received signal of duration $N_s + N'_z$, with $N'_z \leq N_z$, looking for the best trade off between performance and computational complexity.

The impact of the number of guard symbols $N_{z,1}$ in a back-to-back (B2B) scenario is shown in Fig. 4.5(a) for $N_b = 8$, where it can be observed that $N_{z,1} = 30$ is enough (under the other assumptions) to account for the boundary conditions required by the NFT at the TX. A comparison between Fig. 4.5(a) and Fig. 4.5(b), which reports the same for $N_b = 16$, shows that increasing the burst length (and therefore, the energy of the signal), the number of guard symbols for the BNFT increases to around $N_{z,1} = 80$.

While increasing the number of guard symbols $N_{z,1}$, the Q^2-factor saturates at around $\sim 39\,\mathrm{dB}$ at low powers in Figs. 4.5(a)-(b). This saturation is caused by the accuracy NFT operations, determined by the oversampling factors $N_B = N_F = 4$. Indeed, as will be shown in the following, when increasing the oversampling factors, the saturation level increases. Conversely, the required number of guard symbols $N_{z,1}$ increases when increasing the power because the optical signal, obtained through the BNFT, acquires a tail at higher powers, as

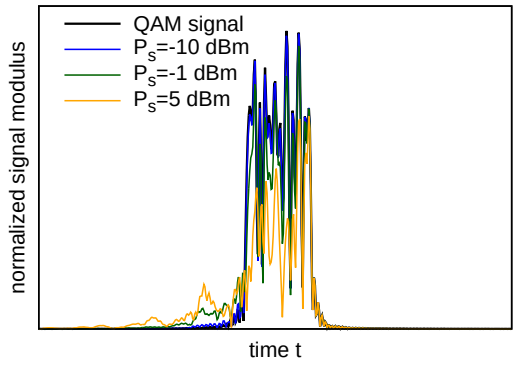


Figure 4.8: Modulus of the initial QAM signal and the optical signals (after BNFT) corresponding to different powers. For the sake of illustration, the signals have been normalized in order to have the same QAM signal. At the lowest power $P_s = -10\,\text{dBm}$, the optical signal is superimposed with the QAM signal, while it acquires a tail (on the left) for increasing power levels.

shown in Fig. 4.8.

The impact of the temporal window considered at the RX for the FNFT processing is shown in Figs. 4.6(a)-(b), for the ideal noise-free and noisy channel, respectively. The figures are obtained for $N_b = 8$ and fixed $N_{z,1} = 30$. Fig. 4.6(a) shows that $N_z' = 700$ is enough to account for GVD at lower powers, but not at higher powers. However, in the noisy case, considering $N_z' = 700$ can also provide a small gain—evident in the optimal power region where the Q-factor is maximum—in terms of performance, as shown in Fig. 4.6(b). This unexpected behavior is due to the fact that, in this case, the processed noisy signal is shorter, and, therefore, the amount of noise entering the FNFT at the RX is smaller.

As already mentioned, performance worsens at lower powers in the noisy scenario in Fig. 4.6(b) with respect to the noise-free scenario in Fig. 4.6(a) because the ASE noise affects the SNR. Conversely, at higher powers, noise-free performance decays because the NFT is a nonlinear operation, and, therefore, one needs to consider higher oversampling factors to achieve the same accuracy. On the other hand, in the noisy scenario in Fig. 4.6(b), performance increases up to an optimal point, as a consequence of the SNR increase, and, then, decays again. As aforementioned, the decay at higher powers may be due to both numerical inaccuracies and the ASE noise. With this in mind, we performed further simulations (in the same scenario) with higher oversampling factors—$N_F = N_B = 40$—shown in Figs. 4.7(a)-(b). Comparison with Figs. 4.6(a)-(b) shows that (i) increasing the oversampling factors improves performance in the noise-free case, as expected, and (ii) in the noisy case performance is not affected by this parameter, implying that they are limited by ASE noise. As a consequence, this reasoning allowed to separate numerical inaccuracies

from physical limitations, and, therefore, to claim that the performance decay in Figs. 4.6(b) and 4.7(b) is caused by the ASE noise and indicates the limitations of the scheme, without being affected by numerical issues. The results shown in this thesis have been obtained paying particular attention to this issue.

4.2.1.2 System performance

Figure 4.9 shows the Q-factor obtained with numerical simulations (solid lines) and with the effective SNR of the analytical model in [64] (dotted lines), for different burst lengths N_b. The corresponding rate efficiencies η are also indicated. The effective SNR in [64], given for Gaussian symbols, is

$$\mathrm{SNR}_{\mathrm{eff}} = \frac{ST_s/N_0}{2(E_{in}/E_{NL})^2 + E_{in}/E_{NL} + 1}, \tag{4.7}$$

where S is the average power per sample of the initial QAM signal (before the NIS), N_0 is the PSD of the accumulated ASE noise, $E_{in} = ST_sN_bN_{ch}$ is the average energy of the initial QAM signal (before the NIS) with N_{ch} being the number of channels ($N_{ch} = 1$ in our case), and E_{NL} is an energy threshold for nonlinear effects $E_{NL} = |\beta_2|N_{ch}/(\gamma T_s)$. If considering QPSK or 16-QAM symbols, Eq. (4.7) can be adjusted adding the term $-c(E_{in}/E_{NL})^2/N_b$ to the denominator in (4.7), with $c = 1$ or $c = 17/25$, respectively. Unless otherwise stated, the simulations are performed with $N_F = N_B = 4$. At high powers, some rare but disruptive numerical instabilities in the calculation of the noisy nonlinear spectrum were observed [78]. We conjecture they are related to the rise of discrete components in the nonlinear spectrum (solitons) induced by noise [13]. This issue has been practically resolved by resorting to linear interpolation between adjacent frequencies when the phenomenon occurs at a given frequency. As explained later, this numerical expedient is not required when employing the windowing technique proposed in Section 4.3. The impact of interpolations is shown in Fig. 4.10(a), which reports the performance obtained with (solid line) and without (dashed line) interpolation.

After reaching a maximum at some optimum power, all curves fall down since the impact of ASE noise on the nonlinear spectrum increases with signal energy (a sort of signal–noise interaction). The impact of numerical inaccuracies, already considered in the previous paragraph, is shown in Fig. 4.10(b), where the simulation results of Fig. 4.9 are compared with the results obtained in the corresponding noise-free scenario and with those obtained with higher sampling rate ($N_B = N_F = 16$ samples per symbol) and longer guard time ($N_z = 900$ guard symbols). It is apparent that, in the region near the optimal power, the noise-free curves are above the noisy ones; moreover, the performance remains unchanged if a higher accuracy is considered. Therefore, Fig. 4.10(b) confirms that the observed performance degradation is due to the interaction of signal and noise, rather than to numerical inaccuracies.

The agreement between theory and simulations in Fig. 4.9 up to the optimum power further confirms that the obtained results are not affected by limitations of the numerical algorithms and that the performance decay at high power is

an intrinsic limitation of this transmission scheme. Moreover, it validates the accuracy of the perturbation approach and asymptotic approximations used in [64] for the computation of the effective SNR.

Figure 4.9 highlights two important facts. Firstly, the maximum Q-factor reduces as N_b increases. This behavior persists for bursts longer than the channel memory, as shown for $N_b = 1024$. Secondly, unlike conventional systems, the performance of the NIS scheme considered here remains unchanged (at least in the considered range of powers) if the optical fiber channel is replaced with an AWGN channel with same accumulated noise (shown with dashed line for $N_b = 32$; results are similar for any N_b). We will return on this later.

To better understand the first issue and its relevance, the maximum of each curve in Fig. 4.9 is reported in Fig. 4.11(a) as a function of the rate efficiency η, and compared with the corresponding performance of a conventional system—i.e., without any NFT—employing ideal DBP as a nonlineairity compensation technique or EDC—i.e., a purely linear technique. For the sake of comparison, burst mode transmission with $N_z = 800$ and same modulation parameters were considered in all cases. As expected, the performance of both DBP and EDC converges to that of a continuous transmission for bursts longer than the channel memory ($N_b > N_z$, corresponding to $\eta > 0.5$). This is because, in these systems, nonlinear interaction involves only signal and noise components that are closer in time than the overall channel memory. On the other hand, the NIS performance keeps decreasing even for longer bursts, as in this case signal–noise interaction does not occur during fiber propagation, as in conventional systems, but at the receiver over the full integration window used for computing the nonlinear spectrum (the FNFT), with an impact that increases with the total signal energy therein. This is confirmed by the curve shown in Fig. 4.9 for the NIS system over an AWGN channel. The comparison in Fig. 4.11(a) also reveals that ideal DBP performs better than NIS, which performs better than EDC only for bursts of short–medium length.

In the single-user scenario considered here, its worse performance with respect to ideal DBP may be not a crucial issue, as NIS is expected to perform better in multi-user scenarios, where DBP is much less effective because of inter-channel interference [12]. A more critical issue is the peculiar dependence of NIS performance on burst length, as confirmed also by theory. In fact, as shown in Fig. 4.11(a), a reasonable performance is obtained only at the expense of a low rate efficiency. For $\eta > 0.11$, NIS performance becomes worse than simple EDC, definitely loosing any appeal. Moreover, also the computational complexity of most practical NFT algorithms has an unfavorable dependence on the total signal length (burst length plus guard time).

Note that, if there were no broadening, a much shorter guard time and burst length could be considered, with a significant improvement in performance and complexity. Thus, one is tempted to check whether it is any better in links with low-dispersion fibers. Considering a sixteen times lower dispersion ($\beta_2 = -1.27\,\mathrm{ps}^2/\mathrm{km}$) and guard time ($N_z = 50$), Fig. 4.11(b) shows that the dependence of NIS performance on the rate efficiency remains practically unchanged (signal–noise interaction in the FNFT is reduced for a shorter burst,

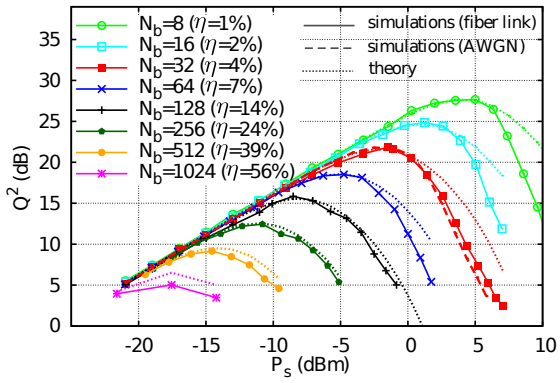


Figure 4.9: Q-factor vs optical launch power for standard NIS with different burst length N_b (and rate efficiency η). QPSK symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 800$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

but is increased for a lower dispersion, the two effects canceling out). Instead, DBP and EDC performance worsens at high rates and slightly improves at low rates. The overall picture does not change significantly: DBP outperforms NIS and EDC, whose performance is almost the same up to $\eta = 0.14$. For $\eta > 0.14$, NIS performance degrades much faster than EDC and DBP, as in the previous scenario.

4.3 Windowing

As mentioned above, time broadening affects also the computational complexity of the FNFT at RX, which, in principle, must be performed for each burst on the entire time range $-\mathcal{T}/2 < t < \mathcal{T}/2$, with $\mathcal{T} = (N_b + N_z)T_s$. Nevertheless, some computational savings can be achieved by noting that, similarly to the linear spectrum, also the continuous part of the nonlinear spectrum experiences a sort of group velocity dispersion during propagation, with different frequency components traveling at different speeds. As a result, different time portions of the received optical signal bring information about different spectral components of the nonlinear spectrum. This is illustrated in Fig. 4.12(a), which shows the modulus of the nonlinear spectrum (vertical axis) as obtained when applying the Boffetta–Osborne algorithm to the received optical signal truncated to the time interval $-\mathcal{T}/2 < t < \tau$, with the upper limit reported on the τ axis. Three different spectral components λ are reported at different depths in the graph. For illustration purposes, results are shown in the absence of optical noise. It is apparent that, for each spectral component λ, only a small portion of the received optical signal—contained in a time window whose center depends linearly on the considered frequency—contributes to the final value of the nonlinear spectrum. This suggests the following windowing technique: given the received optical signal, each frequency component $\rho(L, \lambda)$ is computed by

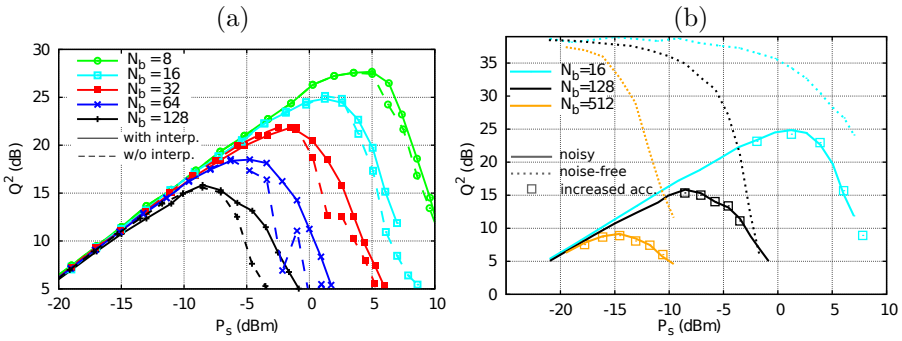


Figure 4.10: Q-factor vs optical launch power for standard NIS (a) with and without interpolation; (b) with original ($N_B = N_F = 4$, $N_z = 800$) and increased ($N_B = N_F = 16$, $N_z = 900$) accuracy, and in the noise-free scenario. QPSK symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 800$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

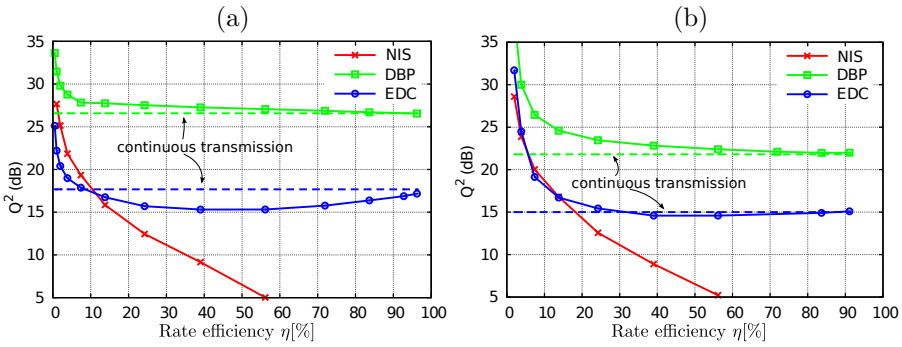


Figure 4.11: Optimal Q-factor vs rate efficiency for NIS and conventional systems with EDC and DBP for (a) same parameters considered in Fig. 4.9 i.e., $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 800$; and (b) $\beta_2 = -1.27\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 50$. QPSK symbols, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

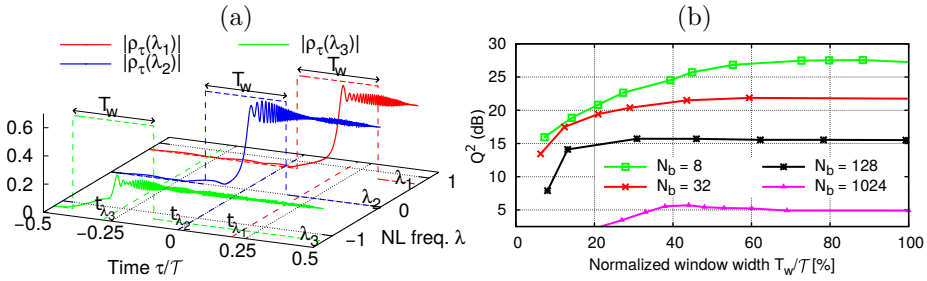


Figure 4.12: (a) Modulus of the noise-free nonlinear spectrum (vertical axis) vs upper time limit considered in the LP algorithm (τ axis) for different spectral components; (b) optimal Q-factor when using the windowing technique vs window width. QPSK symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 800$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

applying the Layer-Peeling (LP) method on the moving time window

$$\max\{t_\lambda - T_w/2, -\mathcal{T}/2\} < t < \min\{t_\lambda + T_w/2, \mathcal{T}/2\}, \tag{4.8}$$

where $T_w < \mathcal{T}$ is the window width (to be optimized) and $t_\lambda = -2\beta_2 L\lambda/T_0$ its center, with T_0 the time normalization parameter.

Fig. 4.12(b) shows the Q-factor at optimum power obtained by the described windowing technique as a function of the window width T_w, for the same system in Fig. 4.9. These results show that the time window for computing the FNFT can be reduced to about 70%, 50%, 20%, and 40% of the total signal duration for $N_b = 8$, 32, 128, and 1024, respectively, with significant computational savings. The different behavior for different burst lengths depends on the maximum achievable Q-factor and on the initial burst length: the lower the Q-factor, the narrower the time window where the signal contribution dominates over noise, until the window width becomes much smaller than $T_s N_b$.

Moreover, this technique avoids the excess noise outside the window of interest for each considered frequency, slightly improving performance (as shown for $N_b = 1024$ in Fig. 4.12(b)) and reducing the numerical instabilities of the LP algorithm mentioned in the previous section. In fact, when using the windowing technique, we were able to reproduce the same results of Fig. 4.9 without the need to resort to the interpolation expedient and almost avoiding the small penalty (compared to the theoretical curves) observed in Fig. 4.9 for $N_b = 1024$ near the optimum launch power.

4.4 Precompensation

The large number of guard spaces N_z required to avoid burst interaction during propagation reduces the transmission rate by the factor $\eta = N_b/(N_z + N_b)$, with a significant loss of spectral efficiency. For mitigating this loss, a precompensation technique allows to reduce N_z, by minimizing the time broadening induced

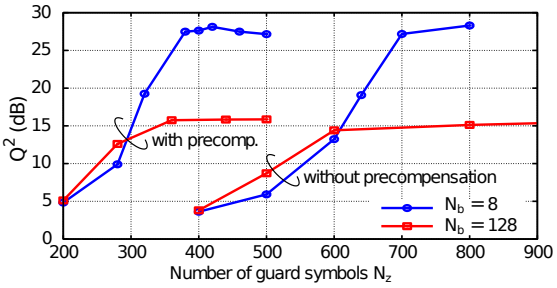


Figure 4.13: Optimal Q-factor vs guard interval with and without precompensation. QPSK symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 800$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

by dispersion throughout the link [75,79,85]. In order to remove the propagation effect from the received nonlinear spectrum $\rho(L,\lambda)$, rather than multiplying it by $e^{j4\lambda^2\mathcal{L}}$ as done in standard NIS, we split the compensation between TX and RX, both pre-multiplying $\rho(0,\lambda)$ and post-multiplying $\rho(L,\lambda)$ by $e^{j2\lambda^2\mathcal{L}}$. This is the same as generating the signal at a distance $-L/2$ and propagating it to a distance $L/2$. In this way, the same time broadening of $\pi|\beta_2|LR_s^2(1+\beta)$ symbol times is observed at TX and RX, in fact halved with respect to the standard implementation.

Fig. 4.13 shows the Q-factor obtained with and without precompensation for bursts of length $N_b = 8$ and $N_b = 128$ at their optimal launch power (about 3.8 dBm and -8.5 dBm, respectively) as a function of the number of guard symbols N_z, with the same simulation setup as in Fig. 4.9. For both burst lengths, precompensation allows using half the guard time to achieve the same performance, with a significant increase of the rate efficiency (almost doubled). This precompensation technique reduces the computational complexity of the FNFT while increasing the BNFT one (as the total processing windows are, respectively, shortened and lengthened), with an overall effect that depends on the algorithms employed for the BNFT and FNFT.

The same technique was independently proposed in [85].

4.5 Detection strategies

NFDM paradigms have been developed borrowing concepts from linear communication and, indeed, can be thought as a nonlinear version of the well known OFDM. However, very little is known about NFT-based transmission schemes and researchers are still working towards its optimization. In particular, the detection strategy commonly considered for NFDM is optimal for conventional systems (assuming AWGN channel), but is not optimal for NFDM since it does not account for the actual statistics of noise in the nonlinear frequency domain—i.e., where the system decides symbols. Indeed, even assuming an

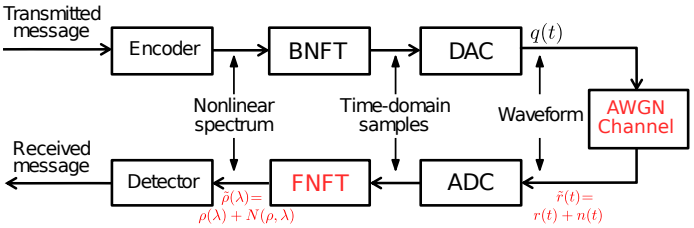


Figure 4.14: NIS scheme with AWGN channel.

AWGN channel (see Fig. 4.14), noise in the nonlinear frequency domain is not AWGN, as the FNFT operation—a nonlinear operation—changes the statistics of noise, which, remarkably, depend on the nonlinear spectrum itself [15, 64]. The considered detection strategy is optimal only at low powers, as in this case the scheme tends to an OFDM scheme. Consequently, the currently achieved NFDM performance can be much improved, and novel detection strategies tailored for NFDM might reveal its potential and allow to actually outperform conventional systems [54, 72, 75–77, 80, 86].

In this Section, after describing a common system setup, three alternative detection strategies are introduced [54, 76, 77]. The decision-feedback BNFT (DF-BNFT) detection, deciding symbols in time domain avoiding any FNFT, provides a significant gain with respect to conventional NFDM, at the expense of a relevant increase of the computational complexity. On the other hand, the incremental FNFT (I-FNFT) and the decision-feedback FNFT (DF-FNFT) detections provide some good performance improvements reducing the noise in the nonlinear frequency domain without, or only slightly, increasing the computational complexity.

In this Section, the BER obtained with the detection strategies P_b is estimated by direct error counting, rather than estimated thought the EVM. Numerical NFT operations are performed with an oversampling factor of $N_F = N_B = 8$ samples per symbols. The FNFT is numerically performed using the Boffetta–Osborne method, while the BNFT is computed with the NCG method, see Section 3.4.

4.5.1 System description

The considered transmission scheme is sketched in Fig. 4.15, and is very similar to the NIS considered in Sec. 4.2 [54, 76, 77].

As in the NIS scheme [50], the TX encodes a burst of N_b symbols $\{x_1, \ldots, x_{N_b}\}$ drawn from an M_X-ary QAM alphabet $\{X_1, \ldots, X_{M_X}\}$ onto a QAM signal (4.1). For reasons that will be clarified later, we restrict the pulse shape $g(t)$ to have a finite duration smaller than T_s. As in NIS, the ordinary FT $S(f)$ of (4.1) is then mapped on the continuous part of the nonlinear spectrum $\rho(\lambda)$ according to Eq. (4.2). Furthermore, before computing the BNFT, deterministic propagation effects (dispersion and nonlinearity) are

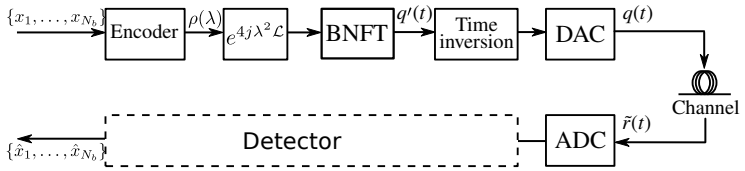


Figure 4.15: NIS system setup for improved detection strategies.

precompensated by multiplying the nonlinear spectrum by $e^{j4\lambda^2\mathcal{L}}$, where $\mathcal{L}$ is the normalized link length. Finally, the input optical signal is taken to be $q(t) = q'(-t)$, where $q'(t)$ is the BNFT of the precompensated nonlinear spectrum.

There are two differences between the TX described here and the one considered in Sec. 4.2, which, however, do not change the overall NIS working principle. Firstly, propagation effects are removed at the TX (precompensation) rather than at the RX. While both solutions are feasible (and splitting the compensation between TX and RX might even be advantageous as described in Sec. 4.4), precompensation is considered here because the proposed strategies relies on it. Indeed, Corollary 13, on which the strategies are based (as explained in the following), applies only to optical signals unaffected by propagation effects. Precompensation ensures that the received signal meet this requirement. Secondly, each burst in the optical signal is inverted in time just to help explaining the working principle, not because it is necessary.

Let us denote by $r(t)$ the optical signal obtained by propagating $q(t)$ in a noise-free channel. In this case, the only channel effect is the multiplication of the nonlinear spectrum by $\exp(-j4\lambda^2\mathcal{L})$, so that $r'(t) = r(-t)$ would be the BNFT of $\rho(\lambda)$, the nonlinear spectrum before precompensation.

Corollary 12. *Considering a NIS modulation, the above optical signal $r(t)$ for $t \leq \tau$ depends only on the values of the QAM signal $s(t)$ for $t \leq \tau$ (here, the variables can be considered as all dimensional or all dimensionless).*

Proof. Let us consider normalized units. Using Eq. (4.3), $F(y)$ for $y \geq 2\tau$ depends only on the values of $s(t)$ for $t \leq -\tau$. Therefore, recalling that $r'(t)$ is the BNFT of $\rho(\lambda)$, the NFT causality property (Proposition 10) implies that $r'(t)$ for $t \geq \tau$ depends only on $s(t)$ for $t \leq -\tau$. Changing the sign of τ, one obtains that $r'(t)$ for $t \geq -\tau$ depends only on $s(t)$ for $t \leq \tau$. Finally, the thesis follows considering that $r(t) = r'(-t)$. The concept can be easily extended to dimensional variables. □

Corollary 13. *Let $t_k = (k - 1/2)T_s$. If the pulse shape $g(t)$ in (4.1) has finite duration $T \leq T_s$, for $t \leq t_k$ the optical signal $r(t)$ depends only on the symbols $x_1, .., x_k$, and not on the following ones $x_{k+1}, .., x_{N_b}$, as shown in Fig. 4.16. In mathematical formulas*

$$r(t)\big|_{t\leq t_k} = \mathcal{G}\left\{x_1, .., x_k\right\}, \tag{4.9}$$

where $\mathcal{G}$ is a generic function.

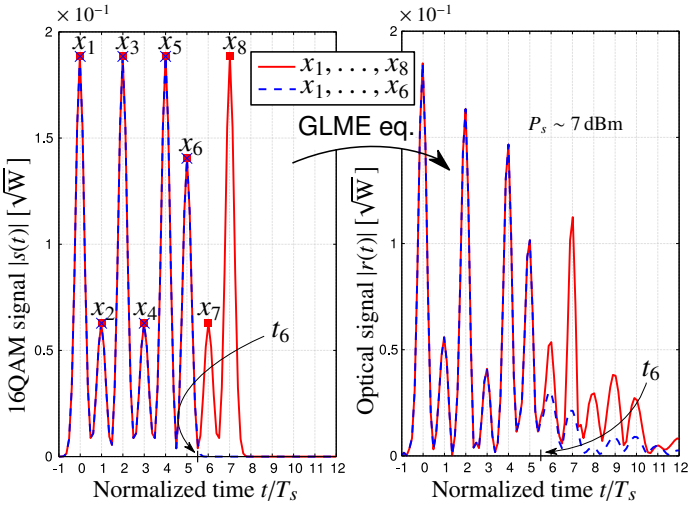


Figure 4.16: The NFT causality property for NIS with no ISI on $s(t)$. A train of Gaussian pulses, modulated by 16QAM symbols, and almost ISI-free, is shown before (on the left) and after (on the right) the BNFT. The red signal is generated by 8 symbols, while for the blue one only the first 6 are taken into account. The two optical signals are superimposed for $t \leq t_6$, as for Eq. (4.9) (baudrate $R_s = 50\,\text{GBd}$, optical power $P_s = 7\,\text{dBm}$).

Proof. With this hypothesis, $s(t)$ depends only on the symbols $x_1, \ldots, x_k$ for $t \leq t_k$. Then, applying Corollary 12, the thesis follows. □

4.5.2 Decision-feedback backward nonlinear Fourier transform detection

The improved detection scheme proposed here originates from the idea that, since a detrimental signal–noise interaction takes place when computing the FNFT of the received noisy signal in NIS, decisions could be alternatively made by comparing the received signal with the BNFT of all possible transmitted (noise-free) waveforms, thus avoiding signal–noise interaction effects. Selecting the waveform (and the corresponding symbols) closest to the received optical signal would correspond to a maximum a posteriori probability (MAP) strategy, under the assumption that the accumulated optical noise can be modeled as AWGN (more on this later). The drawback of a sequence—rather than symbol-by-symbol—detection strategy, is an exponential growth of the detector complexity with the burst length N_b. In order to avoid this growth, the NFT causality property and a decision-feedback scheme are finally employed, obtaining the DF-BNFT detection scheme depicted in Fig. 4.17. Symbols detection in time domain after the FNFT has been considered also in [80] for the detection of 7 eigenvalues, and in [86] using machine learning.

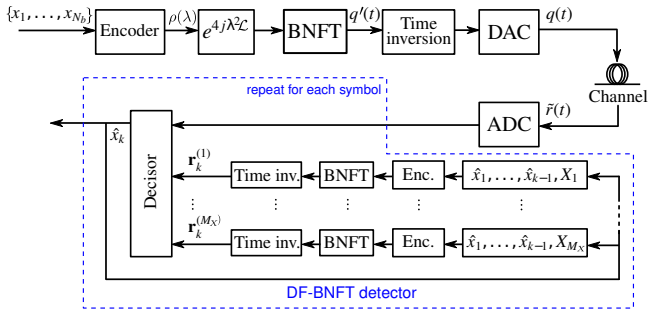


Figure 4.17: NIS with DF-BNFT detection strategy.

Having in mind that the NIS performance decay is caused by the detection strategy itself, rather than signal–noise interaction during propagation—as also hinted by the fact that the performance obtained by simply adding AWGN after noise-free propagation is superimposed with that of the actual optical link, as shown in Fig. 4.9 [75]—the aim of this work is to devise a detection strategy at least optimal for the AWGN channel. Therefore, assuming an AWGN channel, we can write the received noisy optical signal as

$$\tilde{r}(t) = r(t) + n(t), \tag{4.10}$$

where $r(t)$ is, as before, the optical signal obtained by propagating the input signal $q(t)$ in a noise-free channel, and $n(t)$ is circularly-symmetric complex white Gaussian noise with PSD N_0. We would like to stress that (4.10) is only an *ansatz* and not an actual identity.

An ADC recovers the samples of the received noisy optical signal and collects them in the vector $\tilde{\mathbf{r}}$. The ADC is modeled as a rectangular filter with bandwidth $\mu/(2T_s)$ that acquires μ samples per symbol time. Assuming that the filter bandwidth is larger than the overall signal bandwidth, under the AWGN assumption (4.10), $\tilde{\mathbf{r}}$ is a sufficient statistic and we can write $\tilde{\mathbf{r}} = \mathbf{r} + \mathbf{n}$, where $\mathbf{r}$ is a vector collecting the samples of $r(t)$, and $\mathbf{n}$ is a vector of i.i.d. circularly-symmetric complex Gaussian r.v.s n_k, with zero mean and variance $\sigma^2 = E\{|n_k|^2\} = N_0\mu/T_s$. Therefore, conditional on $\mathbf{r}$, the components of $\tilde{\mathbf{r}}$ are independent. For the sake of simplicity, let $\tilde{\mathbf{r}}_k$ (and $\mathbf{r}_k$) be the vector of length ν representing the noisy optical signal (and, respectively, its noise-free equivalent) in the time window $[t_{k-1}, t_k)$. Hence, $\tilde{\mathbf{r}}$ can also be written as a compound vector $\tilde{\mathbf{r}} = (\tilde{\mathbf{r}}_1, \ldots, \tilde{\mathbf{r}}_{N_b})$ containing the samples of the received signal in $[-T_s/2, (N_b - 1/2)T_s)$, i.e., the time window of duration N_bT_s in which information is encoded. Indeed, with respect to the QAM signal $s(t)$, the optical signal after the GLME broadens in time, developing a sort of right tail that extends outside the considered detection window, i.e., for $t > (N_b - 1/2)T_s$, as shown for instance in Fig. 4.16. Therefore, a longer vector $\tilde{\mathbf{r}}$ should be considered to obtain a sufficient statistic. However, in the decision-feedback strategy derived in the following, this tail could be used only to detect the last symbol x_{N_b} of

the sequence, with a negligible contribution to the overall performance. Hence, for the sake of simplicity, we simply discard it.

According to the MAP strategy, and assuming equally likely input symbols, optimal detection maximizes the probability density function (pdf) $p(\tilde{\mathbf{r}}|\mathbf{x})$ of the vector $\tilde{\mathbf{r}}$ conditional upon the transmitted sequence $\mathbf{x} = (x_1, .., x_{N_b})$. Thus, an optimum RX chooses the sequence $\hat{\mathbf{x}}$ according to

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmax}}\, p(\tilde{\mathbf{r}}|\mathbf{x}). \tag{4.11}$$

Since, conditional upon $\mathbf{x}$, the components of $\tilde{\mathbf{r}}$ are independent, the pdf in Eq. (4.11) can be factorized as

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmax}} \prod_{k=1}^{N_b} p(\tilde{\mathbf{r}}_k|\mathbf{x}) = \underset{\mathbf{x}}{\operatorname{argmax}} \sum_{k=1}^{N_b} \ln p(\tilde{\mathbf{r}}_k|\mathbf{x}), \tag{4.12}$$

where the second equality stems from the monotonic behavior of the logarithm.

The NFT property (4.9) implies that the signal samples in the time window $[t_{k-1}, t_k)$, i.e., those collected in $\tilde{\mathbf{r}}_k$, depend only on the symbols $(x_1, \ldots, x_k)$. Therefore, an optimum RX performs decisions according to

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmax}} \sum_{k=1}^{N_b} \ln p(\tilde{\mathbf{r}}_k|(x_1, \ldots, x_k)). \tag{4.13}$$

Since the symbols $x_1, \ldots, x_k$ uniquely determine $\mathbf{r}_k$, under the AWGN assumption we have

$$p(\tilde{\mathbf{r}}_k|(x_1, \ldots, x_k) = \frac{1}{(\pi\sigma^2)^{\mu}} \exp\left(-\|\tilde{\mathbf{r}}_k - \mathbf{r}_k\|^2/\sigma^2\right) \tag{4.14}$$

so that $\ln p\big(\tilde{\mathbf{r}}_k\big|(x_1, \ldots, x_k)\big) = -\mu \ln\left(\pi\sigma^2\right) - \|\tilde{\mathbf{r}}_k - \mathbf{r}_k\|^2/\sigma^2$. This implies that, in order to determine the optimal sequence $\hat{\mathbf{x}}$, all possible $M_X^{N_b}$ input sequences should be considered, which is not a viable solution. In order to avoid this exponential growth of complexity, we resort to a sub-optimal decision-feedback strategy—namely, DF-BNFT—by which symbols are decided iteratively for $k = 1, \ldots, N_b$ as

$$\hat{x}_k = \underset{X_i \in \left\{X_1, .., X_{M_X}\right\}}{\operatorname{argmax}} \ln p\big(\tilde{\mathbf{r}}_k|(\hat{x}_1, ..., \hat{x}_{k-1}, X_i)\big) \tag{4.15}$$

bringing down to $M_X \times N_b$ the number of sequences to be considered. Equation (4.15) can be evaluated by comparing the (samples of the) received signal with (the samples of) M_X trial waveforms $r_k^{(i)}(t)$ uniquely corresponding to the sequences $\hat{x}_1, ..., \hat{x}_k, X_i$ for each X_i in the symbol constellation $\{X_1, .., X_{M_X}\}$. The waveform $r_k^{(i)}(t)$ is obtained from the symbol sequence $\hat{x}_1, ..., \hat{x}_k, X_i$ by the same encoding technique used at the TX, except for precompensation. Let $\mathbf{r}_k^{(i)}$ denote the vector of length μ containing the samples of $r_k^{(i)}(t)$ in $[t_{k-1}, t_k)$, referred to as detection window. In other words, $\mathbf{r}_k^{(i)}$, $i = 1, \ldots, M_X$, are all

possible vectors $\mathbf{r}_k$ given that the sequence $\hat{x}_1, ..., \hat{x}_k, X_i$ has been sent. The RX implements the DF-BNFT strategy through N_b steps as follows.

For $k = 1, \ldots, N_b$:

- Digitally obtain the vectors $\mathbf{r}_k^{(i)}$, for each X_i in the symbol constellation $\{X_1, .., X_{M_X}\}$.
- Choose $\hat{x}_k$ by the rule

$$\hat{x}_k = \underset{X_i \in \{X_1, .., X_{M_X}\}}{\operatorname{argmin}} \left\| \tilde{\mathbf{r}}_k - \mathbf{r}_k^{(i)} \right\|^2 \tag{4.16}$$

  where this last equation follows from (4.14) and (4.15).

The DF-BNFT strategy (4.15) avoids any intersymbol interference (ISI) thanks to decision feedback, which accounts for the dependence of $\tilde{\mathbf{r}}_k$ on previous symbols, and to Corollary 13, which ensures that $\tilde{\mathbf{r}}_k$ does not depend on next symbols. However, it is still suboptimal compared to (4.13) for two reasons. The first reason is that it does not exploit the information about x_k that is contained in the received signal after t_k. In fact, as shown in Fig. 4.16, while in the original QAM signal $s(t)$ (on the left) the information about x_k is fully contained in the time interval $[t_{k-1}, t_k)$, in the corresponding optical signal $r(t)$ (on the right) part of this information goes to times $t > t_k$, the effect becoming more relevant with power and as k increases. An apparent consequence of this effect is that the optical signal $r(t)$ has an average amplitude that decreases with time and a sort of "tail" that extends beyond the duration of the original QAM signal. The second reason is that it is affected by error propagation, as previous decisions $\hat{x}_1, ..., \hat{x}_{k-1}$ might be incorrect, this effect becoming more relevant as k increases, too. Finally, even the strategy (4.13), that is optimal for the AWGN channel (4.10), might be suboptimal on a real fiber link. The effects of error propagation, information loss, and non-AWGN channel statistics are discussed in more detail in Subsection 4.5.2.2.

While the DF-BNFT strategy is conceived to improve system performance (as it will be verified in the following through simulations), it also has some drawbacks compared to the FNFT strategy. Firstly, to fulfill the hypothesis of Corollary 13 and avoid ISI, the pulse shape must be fully confined within a symbol time—a more stringent requirement compared to the conventional Nyquist criterion. This imposes the use of pulses with a wider bandwidth, which has a negative impact on the achievable spectral efficiency. Secondly, full channel precompensation is required to use Corollary 13 at the RX, and, therefore, channel compensation cannot be split between RX and TX to reduce guard intervals [75,85]. Thirdly, as far as it concerns computational complexity, the DF-BNFT scheme requires the evaluation of M_X BNFTs over μN_b points. On the other hand, standard NIS detection computes only one FNFT over a comparable number of points. While an exact comparison between the two strategies depends on the relative complexity and accuracy of the considered BNFT and FNFT algorithms, it is reasonable to assume that the complexity

of the DF-BNFT detector is considerably higher than that of standard FNFT detection. Less complex detection strategies will be discussed later in Section 4.5.3.

4.5.2.1 Error probability estimation and bounds

For a given sequence, the probability of error P_e of the DF-BNFT strategy can be evaluated by averaging (over all constellation symbols and symbols within a burst) the probability $P_k^{(m)}$ that an error occurs when the symbol X_m is sent at position k, provided that the symbols $\hat{x}_1, ..\hat{x}_{k-1}$ are correctly detected. Specifically,

$$P_e = \frac{1}{M_X N_b} \sum_{k=1}^{N_b} \sum_{m=1}^{M_X} P_k^{(m)}. \tag{4.17}$$

As computing $P_k^{(m)}$ may be difficult, we will bound it through standard techniques [39]. By defining the events

$$E_{m,i} = \{X_i \text{ is preferred to } X_m \text{ when deciding on } x_k\} \tag{4.18}$$

and

$$E_m = \{X_m \text{ is not preferred when deciding on } x_k\} = \bigcup_{\substack{i=1 \\ i\neq m}}^{M_X} E_{m,i} \tag{4.19}$$

we can upper bound $P_k^{(m)}$ by the union bound

$$\begin{aligned} P_k^{(m)} = P\left(E_m \;\middle|\; (\hat{x}_1, ..., \hat{x}_{k-1}, X_m)\right) &= P\left(\bigcup_{\substack{i=1 \\ i\neq m}}^{M_X} E_{m,i} \;\middle|\; (\hat{x}_1, ..., \hat{x}_{k-1}, X_m)\right) \\ &\leq \sum_{\substack{i=1 \\ i\neq m}}^{M_X} P\left(E_{m,i} \;\middle|\; (\hat{x}_1, ..., \hat{x}_{k-1}, X_m)\right) \end{aligned} \tag{4.20}$$

Due to our AWGN assumption, the pairwise error probabilities in (4.20) are given by

$$P\left(E_{m,i} \;\middle|\; (\hat{x}_1, ..., \hat{x}_{k-1}, X_m)\right) = \mathcal{Q}\left(\frac{d_k^{(m,i)}}{2\sigma}\right) \tag{4.21}$$

where

$$d_k^{(m,i)} = \left\|\mathbf{r}_k^{(m)} - \mathbf{r}_k^{(i)}\right\| \tag{4.22}$$

is the Euclidean distance between $\mathbf{r}_k^{(m)}$ and $\mathbf{r}_k^{(i)}$, and $\mathcal{Q}(x)$ is the Q-function (2.15). We can also obtain a useful approximation on $P_k^{(m)}$ as follows. Denoting by $C_{m,i}$ the event complementary to $E_{m,i}$, we have

$$\begin{aligned} P_k^{(m)} = \;& P\left(\bigcup_{\substack{i=1 \\ i\neq m}}^{M_X} E_{m,i} \;\middle|\; (\hat{x}_1, ..., \hat{x}_{k-1}, X_m)\right) = \\ & = 1 - P\left(\bigcap_{\substack{i=1 \\ i\neq m}}^{M_X} C_{m,i} \;\middle|\; (\hat{x}_1, ..., \hat{x}_{k-1}, X_m)\right) \end{aligned} \tag{4.23}$$

and, taking into account that $P\big(C_{m,i} \mid (\hat{x}_1, ..., \hat{x}_{k-1}, X_m)\big) = 1 - P\big(E_{m,i} \mid (\hat{x}_1, ..., \hat{x}_{k-1}, X_m)\big)$,

$$P_k^{(m)} \simeq 1 - \prod_{\substack{i=1 \\ i \neq m}}^{M_X} \Big(1 - P\left(E_{m,i} \mid (\hat{x}_1, ..., \hat{x}_{k-1}, X_m)\right)\Big) \tag{4.24}$$

where the approximation is due to the fact that, in general, the events $C_{m,i}$ are not mutually independent. Let us now derive a lower bound. Recalling that the probability of a union of events is lower bounded by each one of the probabilities of the single events, we have

$$\begin{aligned} P_k^{(m)} &= P\left(\bigcup_{\substack{i=1 \\ i \neq m}}^{M_X} E_{m,i} \,\middle|\, (\hat{x}_1, ..., \hat{x}_{k-1}, X_m)\right) \\ &\geq \max_{i \neq m} P\left(E_{m,i} \mid (\hat{x}_1, ..., \hat{x}_{k-1}, X_m)\right). \end{aligned} \tag{4.25}$$

In conclusion, from (4.20), (4.24), (4.25), and taking into account (4.21), we have the following upper bound, approximation and lower bound, respectively, on $P_k^{(m)}$

$$P_k^{(m)} \leq \sum_{\substack{i=1 \\ i \neq m}}^{M_X} \mathcal{Q}\left(\frac{d_k^{(m,i)}}{2\sigma}\right) \tag{4.26}$$

$$P_k^{(m)} \simeq 1 - \prod_{\substack{i=1 \\ i \neq m}}^{M_X} \left(1 - \mathcal{Q}\left(\frac{d_k^{(m,i)}}{2\sigma}\right)\right) \tag{4.27}$$

$$P_k^{(m)} \geq \mathcal{Q}\left(\frac{d_k}{2\sigma}\right) \tag{4.28}$$

where

$$d_k = \min_{i \neq m} d_k^{(m,i)} \tag{4.29}$$

and $d_k^{(m,i)}$ is as in (4.22). Replacing (4.26)–(4.28) into (4.17) gives the corresponding bounds and approximation on P_e. We remark that both the bounds and the approximation were derived by assuming that the channel is AWGN and that previous decisions are correct. Deviations from this ideal situation might invalidate the bounds and slightly reduce the accuracy of the approximation, as shown in Subsection 4.5.2.3 by numerical simulations.

The above estimates of the probability of error for a given sequence should be averaged over all possible $M_X^{N_b}$ sequences to obtain the average error probability. However, to speed up computation, one can think of averaging over randomly generated sequences until the result stabilizes. As we will show in Subsection 4.5.2.3, this practical approach still provides a reasonable accuracy and a significant computational saving compared to direct error counting.

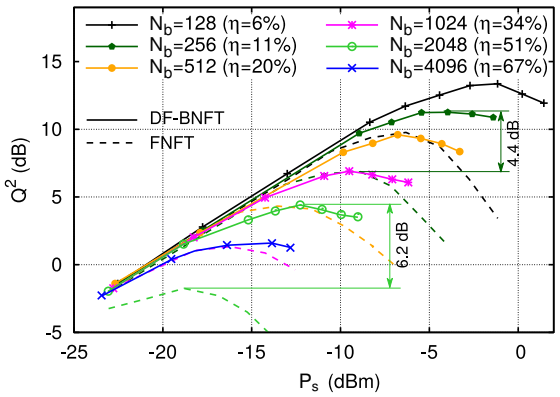


Figure 4.18: Performance of the NFDM system for DF-BNFT (solid lines) and standard FNFT (dashed lines) detection for different burst length N_b (and rate efficiency η). 16QAM symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 2000$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

4.5.2.2 System performance

We simulated the system described in the previous sections and sketched in Fig. 4.17 by using a 16QAM signal $s(t)$ with symbol rate $R_s = 1/T_s = 50\,\mathrm{GBd}$. To fulfill the no-ISI requirement on $s(t)$, the supporting pulse was chosen to be $g(t) = \exp\big(-12.5(t/T_s)^2\big)$, i.e., a Gaussian pulse with a full width at half maximum (FWHM) of $(2/5)\sqrt{2\ln 2}\,T_s \simeq T_s/2$, so that about 99.9% of the energy of a pulse is contained in a symbol time. This pulse shape is chosen to fulfill the requirement of Corollary 13 on the duration of $g(t)$ with only a moderate bandwidth increase compared to the root-raised-cosine pulse employed in Section 4.2 [75]. Note that, by relaxing the energy constraint, the bandwidth could be further reduced to approach the one in Section 4.2. This, however, would also reduce the performance due to the ISI generated by the pulse tail, with an overall effect on the achievable spectral efficiency that should be carefully considered. A typical example of generated signal is shown in Fig. 4.16 on the left.

The channel, otherwise differently stated, is the same considered in Section 4.2, and the bandwidth of both the DAC and the ADC is 100 GHz. The simulation results are deemed free from numerical inaccuracies, since it was verified that the noise-free performance was sufficiently higher than the noisy one (when applicable), and that increasing the accuracy of the numerical algorithms does not change the results, similarly to what done in [78] and explained in Section 3.4.

Figure 4.18 shows the NFDM performance obtained with DF-BNFT detection (solid lines), and with conventional FNFT detection (dashed lines), for different burst lengths (same color for same length). As can be seen, the performance obtained with DF-BNFT is significantly better than that obtained

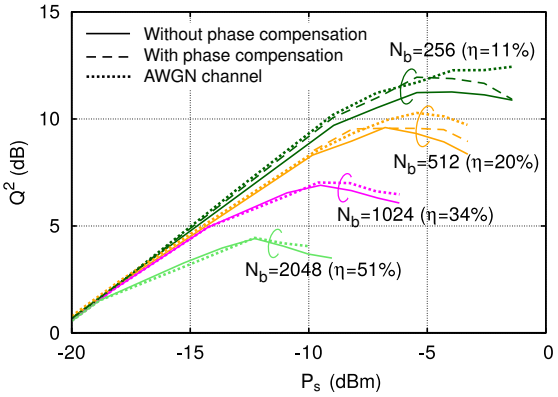


Figure 4.19: Impact of fiber propagation: performance of DF-BNFT on the fiber link without (solid lines) and with average nonlinear phase compensation (dashed lines) and on the AWGN channel (dotted line). Same scenario of Fig. 4.18. 16QAM symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 2000$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

with FNFT detection, with an improvement of 4.4 dB for $N_b = 256$ and 6.2 dB for $N_b = 2048$. However, performance still decays when increasing N_b. This behavior may be due either to the suboptimality of the DF-BNFT detection, or to an intrinsic limitation of the NIS modulation format. For what concerns suboptimality, there are three possible causes of performance degradation, already discussed at the beginning of this Section: the non-AWGN statistics of the fiber channel, which is affected by signal–noise interaction; the error propagation in the decision-feedback mechanism of (4.15); and the information loss entailed by (4.16), which neglects the information about x_k that is contained in the received signal for $t > t_k$. All these effects become more relevant as the burst length increases. Also signal–noise interaction increases with the burst length, as the optical noise interacts with a longer portion of non-zero signal. This effect, however, saturates when the burst length becomes longer than the channel memory.

One may wonder whether the NIS performance shown here is in accordance with the theoretical estimation of the SNR given in [64]. Such a comparison is shown in Fig. 4.9 for almost the same system configuration considered here. The only differences are: (1) the modulation format, which however does not significantly change the results, and (2) the chosen pulse shape. In this work, a pulse with a shorter time duration and, hence, a wider bandwidth is considered. This difference is responsible for a slight performance improvement.

The impact of signal–noise interaction during propagation can be estimated by comparing the performance obtained on the fiber link with the performance obtained on the corresponding AWGN channel (i.e., without fiber propagation and with the same amount of accumulated ASE noise), shown in Fig. 4.19 with dotted lines for $N_b = 256$, 512, 1024, 2048. For longer bursts, i.e., $N_b \geq 1024$,

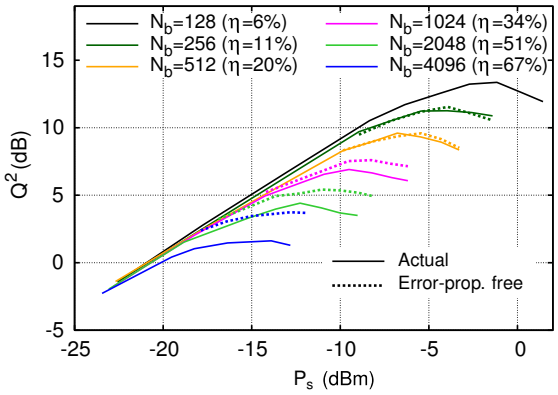


Figure 4.20: Impact of error propagation due to decision feedback in the proposed DF-BNFT detection strategy: actual system performance (solid lines), and error-propagation-free performance (dotted lines). Same scenario of Fig. 4.18. 16QAM symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 2000$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

the corresponding dotted and solid lines are almost indistinguishable. On the other hand, for shorter bursts, i.e., $N_b = 256$, 512, the slight difference between the dotted and solid lines denotes a small impact of signal–noise interaction on system performance and a slight deviation of channel statistics from the AWGN assumption (4.10). One of the effects of signal–noise interaction during propagation is a constant phase rotation of the optical signal. This deviation can be estimated and removed from the optical signal by considering for detection $\tilde{r}(t)e^{-j\alpha}$, α being the phase shift, rather than $\tilde{r}(t)$ itself. Fig. 4.19 shows that a small performance gain can be obtained with this technique (the performance is shown with dashed lines) and that the performance approaches that of the equivalent AWGN channel. Obviously, when the performance is already superimposed to that of the AWGN channel, this technique does not affect the detection strategy.

As regards error propagation, its impact can be estimated from Fig. 4.20, in which the actual DF-BNFT performance (the performance shown in Fig. 4.18) is compared to that of an ideal detector that makes decisions according to the same strategy (4.15), but using the correct symbols $x_1, \ldots, x_{k-1}$ rather than the detected ones $\hat{x}_1, \ldots, \hat{x}_{k-1}$. Contrarily to what observed for signal–noise interaction, the impact of error propagation is more relevant for longer bursts, while it tends to be negligible for shorter ones. Indeed, for longer bursts, farther symbols *(i)* affect more significantly detection, and *(ii)* are more likely to be wrong, since the probability of error is higher.

As regards the third possible cause of performance degradation, further investigations are required to estimate the impact of information loss due to the suboptimality of (4.15) and to devise a better strategy to avoid it. Eventually, the implementation of an optimal strategy based on (4.13), but with a feasible

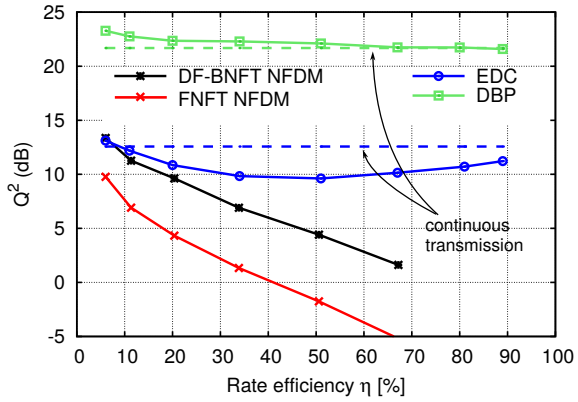


Figure 4.21: Best achievable performance vs rate efficiency for NFDM with different detection strategies and for conventional systems with EDC or DBP. Same scenario of Fig. 4.18, i.e., 16QAM symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 2000$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

complexity, would allow to estimate the ultimate performance of NIS modulation and to understand if the observed performance decay is due to suboptimal detection or to an intrinsic limitation of this modulation scheme. This will be investigated in the future.

The maximum performance achieved by FNFT and DF-BNFT detection (at their respective optimum launch power) in Fig. 4.18 are reported in Fig. 4.21 as a function of the rate efficiency and compared with the maximum performance achieved by conventional systems (also operating in burst mode, for a fair comparison) employing ideal EDC and DBP without any NFT (practically implemented by the split-step Fourier method with 100 steps per span of fiber, enough to practically achieve a perfect compensation of deterministic nonlinearity). DBP performance is estimated from the error vector magnitude [38], rather than calculated by direct error counting, being the corresponding error probability too low to be measured. The improvement of DF-BNFT with respect to FNFT is quite relevant and slightly increases with the rate efficiency η. However, the DF-BNFT performance is still not on par with that of conventional systems—EDC and DBP—and keeps decreasing at higher rates, when the performance of conventional systems saturate to the one achieved with continuous (non burst-mode) transmission. Finally, the performance achievable by DF-BNFT detection was investigated also in different scenarios. Fig. 4.22(a) refers to the same system setup used in Figs. 4.18–4.21 but with a lower dispersion parameter $\beta_2 = -1.27\,\mathrm{ps}^2/\mathrm{km}$ and, therefore, a lower number of guard symbols $N_z = 125$. The overall results do not change significantly, as already observed in [75], and DF-BNFT achieves a performance improvement of almost 6 dB with respect to FNFT detection. The behavior also does not change when using QPSK symbols in the otherwise same system of Fig. 4.18 but with lower

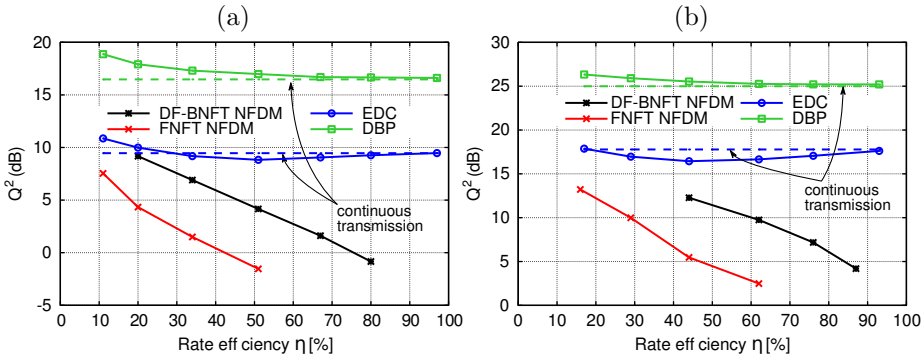


Figure 4.22: Best achievable performance vs rate efficiency for NFDM with different detection strategies and for conventional systems with EDC or DBP: (a) low-dispersion fiber with 16QAM symbols, $\beta_2 = -1.27\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 125$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$; (b) QPSK symbols with $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 160$, $L = 4000\,\mathrm{km}$, and $R_s = 10\,\mathrm{GBd}$.

symbol rate $R_s = 10\,\mathrm{GBd}$, longer link length $L = 4000\,\mathrm{km}$, and $N_z = 160$ guard symbols, as shown in Fig. 4.22(b).

It is worth noting that the position of FNFT detection curve versus those of conventional systems seems worse in Figs. 4.21, 4.22(a)-(b) with respect to the Figs. 4.11(a)-(b). This is due to the fact that, since we are using a different pulse shape, the guard times are much longer and, therefore, the rate efficiency values get smaller. For example, comparing Fig. 4.11(b) and Figs. 4.22(a) (both corresponding to QPSK symbols with $R_s = 50\,\mathrm{GBaud}$, $L = 2000\,\mathrm{km}$, and $\beta_2 = -1.27\,\mathrm{ps}^2/\mathrm{km}$), one should note that the value 56% for the rate efficiency in Fig. 4.11(b), which is evaluated with $N_b = 64$ and $N_z = 50$, corresponds to the value 34% in Figs. 4.22(a), since $N_z = 125$ in the latter case.

4.5.2.3 Validation of the approximation and bounds

The bounds and approximation obtained by replacing (4.26)–(4.28) into (4.17) are reported in Fig. 4.23 for $N_b = 256$ and $N_b = 1024$, after conversion to Q-factor according to (2.20). As the Q-factor is directly related to the bit-error probability P_b, the approximation $P_b \simeq P_e/M_X$ is used, assuming that a symbol error always corresponds to a single bit error. Moreover, if P_e increases, the Q factor decreases, and the other way around. Therefore a lower (or upper) bound for P_e becomes an upper (or lower) bound for Q^2_{dB}. In order to check their accuracy, they are compared with the performance obtained by numerical simulations and direct error counting for the actual fiber channel. In both cases, the approximation lies between the bounds, asymptotically approaching the lower bound when power increases. At low powers, the approximation is in very good agreement with numerical simulations. On the other hand, near the optimum power, the approximation overestimates the actual performance, which falls

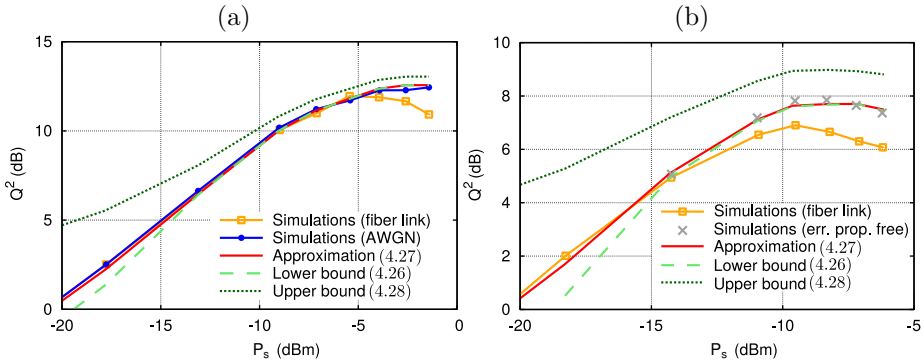


Figure 4.23: Validation of the semianalytic approximation and bounds for the performance of DF-BNFT detection. Same scenario as Fig. 4.18, with (a) $N_b = 256$ and (b) $N_b = 1024$. 16QAM symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 2000$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

slightly below the lower bound. This is due to signal–noise interaction during propagation (for $N_b = 256$) and to error propagation in the decision-feedback strategy (for $N_b = 1024$), both neglected in the derivation of the bounds and approximation (4.26)–(4.28). In fact, when considering the numerical simulations for the AWGN channel in Fig. 4.23a, and the error-propagation-free simulations in Fig. 4.23b, they correctly fall between the bounds and are in excellent agreement with the approximation.

As already explained, the probability of error for a given sequence P_e in (4.17) should be averaged over all possible sequences. However, the number of possible sequences $M_X^{N_b}$ is practically unmanageable, making it impossible performing an exact average. Anyway, most sequences contribute in the same way to the average, so that we don't need to explore all of them, but only account for the most significant ones. This can be done by performing a Monte Carlo average, consisting in randomly generating sequences until the corresponding average performance stabilizes. This is in contrast with the full numerical estimation used in Section 4.5.2.2, in which also the effect of noise is numerically estimated by averaging over many random realizations. To illustrate the difference between the two approaches and show the speed of convergence of the various estimates, Fig. 4.24 reports the same bounds and estimates shown in Fig. 4.23a as a function of the number of iterations (corresponding to the number of sequences of length N_b over which the performance is averaged), considering two different values of the mean power. As can be seen, the computation of the semianalytical bounds and approximation requires only a few iterations in all cases, while computing the performance by full numerical simulations requires a number of iterations that depends on the actual error probability and, hence, on the input power, such that a sufficient number of error events are observed. As an example, $30 \div 40$ iterations suffice for an input power of $-9\,\mathrm{dBm}$, while more than 200 iterations are necessary at the optimum input power of $-4\,\mathrm{dBm}$.

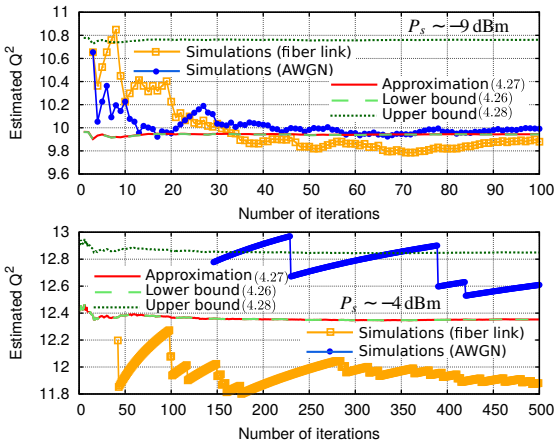


Figure 4.24: Convergence of the numerical simulations and of the semianalytic approximation and bounds with the number of iterations (transmitted sequences). Same scenario as Fig. 4.18, with $N_b = 256$ at $P_s = -9$dBm (above) and at optimal power $P_s = -4$dBm (below). 16QAM symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 2000$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

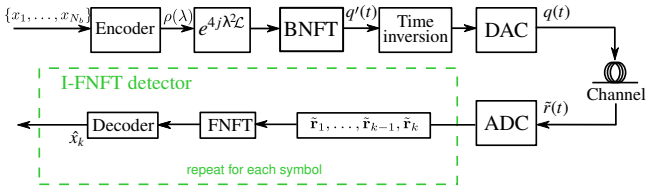


Figure 4.25: NIS with the I-FNFT detection strategy.

4.5.3 Decision-feedback forward nonlinear Fourier transform and incremental forward nonlinear Fourier transform detections

This Subsection presents two detection strategies based on the NFT causality property 10 and Corollary 13 that, similarly to standard detection and differently from the DF-BNFT detection, take decision in the nonlinear frequency domain, but *cleaning* the signal from excessive noise and unessential signal. In the following description we use the same notation as in the previous subsection i.e., $\tilde{\mathbf{r}}_k$ (and $\mathbf{r}_k$) indicates the vector of length μ containing the samples of the noisy (and noise-free equivalent) optical signal in the time interval $[t_{k-1}, t_k)$. Finally, let $r^{(k)}(t) = r(t)|_{t<t_k}$ and $\tilde{r}^{(k)}(t) = \tilde{r}(t)|_{t<t_k}$.

The I-FNFT strategy, sketched in Fig. 4.25, decides symbols in an iterative way in the nonlinear frequency domain, analyzing only a portion of the received signal to reduce noise in the nonlinear spectrum, which increases with signal energy [64]. Specifically, the k-th symbol is detected operating as in conventional

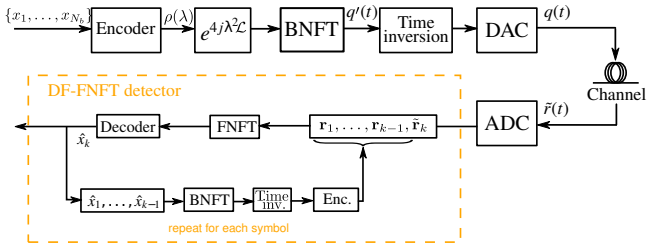


Figure 4.26: NIS with the DF-FNFT detection strategy.

FNFT detection (i.e., FNFT, matched filtering, and sampling), on the samples sequence

$$\tilde{\mathbf{r}}_1,\ldots,\tilde{\mathbf{r}}_{k-1},\tilde{\mathbf{r}}_k \tag{4.30}$$

which represents the continuous signal

$$\tilde{r}_k(t)=\begin{cases}\tilde{r}(t) & t<t_k\\ 0 & \text{else}\end{cases}. \tag{4.31}$$

This detection strategy can be implemented with the same computational complexity as the standard FNFT detection. Indeed, the nonlinear spectrum is computed (e.g., with the Boffetta–Osborne method [7,11]) by recursively adding a small portion of the optical signal and multiplying for the transfer matrix of this contribution. In other words, at the k-th step, one has already computed the contribution of the optical signal for $t<t_{k-1}$ and needs to add only the contribution of the signal in $[t_{k-1},t_k)$, i.e., $\tilde{\mathbf{r}}_k$, resulting overall in a single FNFT.

The DF-FNFT strategy adds a further step to the I-FNFT one: besides considering only a portion of the signal in detection, it also takes advantage of the feedback given by already decided symbols to *clean* the received signal. Specifically, given the symbols $\hat{x}_1,\ldots,\hat{x}_{k-1}$ already decided, the k-th symbol is decided with two steps: (i) digitally perform a BNFT to obtain for $t<t_{k-1}$ the samples of the noise-free signal $r^{(k-1)}(t)$ corresponding to the symbol sequence $\hat{x}_1,\ldots,\hat{x}_{k-1}$ and generates the sample sequence $\mathbf{r}_1,\ldots,\mathbf{r}_{k-1}$ (this is obtained performing the same operation of the TX, but for precompensation), and (ii) perform standard detection (i.e., FNFT, matched filter, and sampling) on the samples

$$\mathbf{r}_1,\ldots,\mathbf{r}_{k-1},\tilde{\mathbf{r}}_k \tag{4.32}$$

which correspond to the continuous signal

$$\tilde{r}_k(t)=\begin{cases}r_{k-1}(t) & t<t_{k-1}\\ \tilde{r}(t) & t_{k-1}\le t<t_k\\ 0 & \text{else}\end{cases}\quad\text{with } r_1(t)=\tilde{r}(t) \tag{4.33}$$

to detect $\hat{x}_k$. DF-FNFT requires to perform at the RX a total of one BNFT and two FNFTs. Indeed, as far as it concerns (i), at the k-th step one needs

to evaluate $r^{(k-1)}(t)$ by performing a BNFT only for $t \in [t_{k-2}, t_{k-1})$, since the values for $t < t_{k-2}$ have already been evaluated at the previous step, resulting overall in a single BNFT. Regarding (ii), similarly to the I-FNFT case, one needs to add the contribution of the signal in two symbol times for $t \in [t_{k-2}, t_k)$, therefore resulting in two FNFTs.

Remarkably, both detection strategies, as well as the DF-BNFT one, choose the k-th symbol x_k accounting only for its contribution in the time window $[t_{k-1}, t_k)$. While x_k does not contribute to the signal before t_{k-1}, it does for $t \geq t_k$, with this contribution increasing at higher energies. Therefore, these detection strategies do not consider all the available information, thus reducing the effective SNR. However, removing part of the signal also improves performance, as shown in the following. Moreover, for what concerns DF-FNFT, considering $r^{(k)}(t) = \tilde{r}(t)$ for $t \geq t_{k-1}$ (i.e., without removing part of signal) would drive to a much more computationally complex detection: at the k-th step, one should perform FNFT adding the contribution of the signal to $N_b - k + 2$ symbols (rather than 1).

System performance System performance was evaluated through simulations, in the same system setup considered in Fig. 4.22(b).

The performance obtained through simulations with FNFT, I-FNFT, DF-FNFT, and DF-BNFT detection are shown in Figs. 4.27(a), 4.27(b), and 4.28(a) for $N_b = 128, 256, 512$, respectively. Firstly, the figures show that the standard FNFT detection for NFDM has the worst performance, as a consequence of being a detection strategy not optimal in the nonlinear frequency domain. Secondly, I-FNFT performs better than FNFT detection allowing for an improvement of up to 3 dB without increasing the computational complexity at all. Next, the figures show that a further performance improvement can be achieved with DF-FNFT detection, at the expense of increasing the computational complexity by additionally performing one FNFT and one BNFT. Finally, DF-BNFT detection provides the best performance, with a gain of up to about 3 dB with respect to DF-FNFT, and about 7 dB with respect to the conventional FNFT.

Figure 4.28(b) reports the optimal performance as a function of the rate efficiency η. The figure shows that increasing the rate efficiency, i.e., the number of information symbols per burst, performance decreases [54,64]. Moreover, Fig. 4.28(a) emphasizes the relative behavior of the considered detection strategies: FNFT performs worse than all others, I-FNFT achieves better results than FNFT, but worse than DF-FNFT, DF-BNFT performs better than all others.

4.6 Exponential mapping and b-modulation

The previous sections of this Chapter were focused on NIS modulation. This Section considers other recently proposed paradigms for modulating information on the optical signal, using the NFT. The results reported in this Section

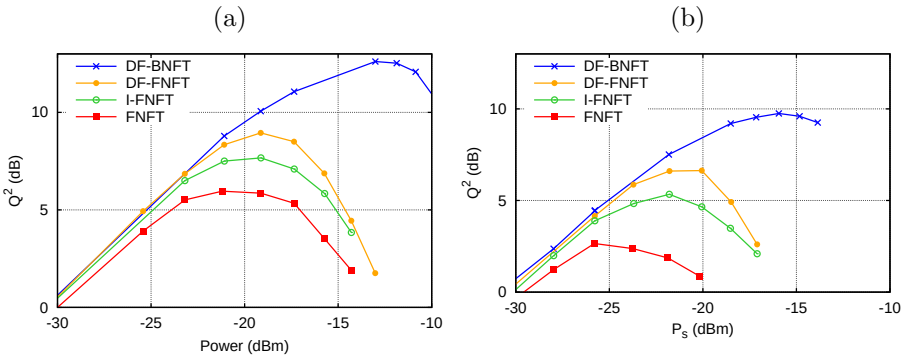


Figure 4.27: NFDM performance for different detection strategies vs launch power for: (a) $N_b = 128$ ($\eta = 44\%$), and (b) $N_b = 256$ ($\eta = 62\%$). QPSK symbols with $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 160$, $L = 4000\,\mathrm{km}$, and $R_s = 10\,\mathrm{GBd}$.

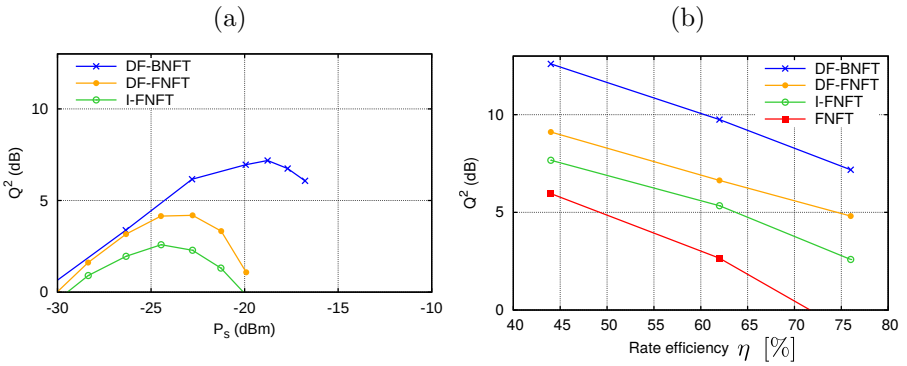


Figure 4.28: NFDM performance for different detection strategies: (a) vs launch power for $N_b = 512$ ($\eta = 76\%$), and (b) optimal performance as a function of the rate efficiency η. QPSK symbols with $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 160$, $L = 4000\,\mathrm{km}$, and $R_s = 10\,\mathrm{GBd}$.

have been obtained in the same scenario as Fig. 4.9 but using the Frumin algorithm [66] rather than the NCG method for the computation of the BNFT; this, however, does not affect the results.

Exponential mapping The exponential mapping has been firstly proposed in [12], and was later considered also in [49, 87] for dual-polarization schemes. Similarly to NIS, the information is mapped on $s(t)$, and its FT $S(f)$ is considered. Next, differently from NIS, the continuous nonlinear spectrum is defined as

$$\rho(\lambda) = \sqrt{\sigma(-1 + e^{\sigma|S(\lambda/\pi)|^2})} e^{j\mathrm{Arg}(S(\lambda/\pi))}, \tag{4.34}$$

and the optical signal is obtained as $q(t) = \mathrm{BNFT}[\rho(\lambda)]$. It follows that the phase of the two spectra—the linear spectrum $S(f)$ and the nonlinear one $\rho(\lambda)$—is the same, while their modulus is different. Moreover, it can be verified that using (4.34) $|\rho(\lambda)| \geq 0$ if $\sigma = 1$, and $0 \leq |\rho(\lambda)| \leq 1$ if $\sigma = -1$. The frequencies f are mapped on the nonlinear frequencies $\lambda = \pi f$. At the RX, the inverse operation is performed[2]

$$S(f) = \sqrt{\sigma \log(1 + \sigma|\rho(\pi f)|^2)} e^{j\mathrm{Arg}(\rho(\pi f))}. \tag{4.35}$$

There are two main reasons for considering this mapping:

1. In the defocusing regime $\sigma = -1$ the nonlinear spectrum must satisfies the following inequality
$$|\rho(\lambda)| \leq 1 \tag{4.36}$$
as follows from Proposition 9; the mapping (4.34) respects this requirement.

2. The mapping (4.34) seems an appropriate choice motivated by Proposition 8. Indeed, the exponential mapping implies that the energy of the QAM signal $s(t)$ and of the optical signal $q(t)$ is the same.

Moreover, similarly to NIS, at low powers,

$$q(t) \to -\sigma s^*(t), \tag{4.37}$$

i.e., at low powers the modulus of the optical signal tends to that of the QAM signal[3]. While the optical signal generated with NIS mapping shows a typical tail (as shown also in Fig. 4.16), the exponential mapping causes tails on both sides of the optical signal, as shown in Fig. 4.29[4], which compares the optical signals (same optical power, but different power for $s(t)$) obtained with the two

[2]The difference in mapping with Eqs. (42)-(47) of [12] is due to the different FT definition considered. In this thesis, the FT is defined according to Eq. (2.11).

[3]Indeed, at low powers $\rho(\lambda) \to \sqrt{|S(\lambda/\pi)|^2} e^{j\mathrm{Arg}(S(\lambda/\pi))} = S(\lambda/\pi)$ from (4.34), and $q(t) \to -\sigma/\pi \left[\int_{-\infty}^{+\infty} \rho(\lambda) e^{j2t\lambda} d\lambda\right]^*$ from (3.65). Combining the two one obtains that $q(t) \to -\sigma s^*(t)$.

[4]For the ease of comparison with Fig. 4.16 and Fig. 4.30 the optical signal shown is $q(t) = q'(-t)$ with $q'(t)$ obtained through NFT similarly to Section 4.5.

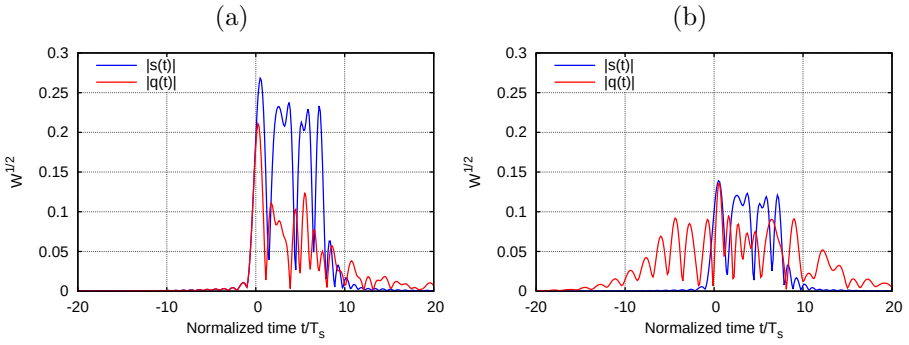


Figure 4.29: Modulus of the optical signal and the QAM signal for with (a) NIS mapping; and (b) exponential mapping. (baudrate $R_s = 50\,\text{GBd}$, optical power $P_s = 10\,\text{dBm}$, QPSK symbols modulated with RRC pulse shape as in Fig. (4.9)).

techniques and the corresponding QAM signal. Unfortunately, Corollary 13 does not hold with this mapping, as shown with a counterexample in Fig. 4.30, and therefore, the detection strategies proposed in Section 4.5 can not be used. Furthermore, it is worth noting that the optical signal obtained with exponential mapping has more energy outside the domain in which $s(t)$ is defined.

Figure 4.31 compares the performance obtained with exponential mapping with those obtained with NIS in the same scenario as in Fig. 4.9. The figure shows that (i) for $N_b = 8$ NIS performs slightly better than the exponential mapping, (ii) for $N_b = 64$ they performs the same, and (iii) for $N_b = 256$ NIS performs slightly worse. However, the RX intensely used interpolation after the evaluation of $a(\lambda)$ (see Section 4.2) and performance might be affected by this issue (increasing the oversampling factors does not improve performance). We will return on this later.

b-modulation In a nutshell, b-modulation consists in encoding the information directly on the scattering data $b(\lambda)$ and $\{\lambda_i, b(\lambda_i)\}_{i=1}^{N_{\text{DS}}}$ rather than on the nonlinear spectrum $\rho(\lambda) = b(\lambda)/a(\lambda)$ and $\{\lambda_i, C_i = b(\lambda_i)/a'(\lambda_i)\}_{i=1}^{N_{\text{DS}}}$. The b-modulation was initially proposed in [72] to reduce the impact of noise at the RX for the modulation of discrete eigenvalues, and has been used for the same purpose in [26,68,88]. Next, the author of [89] proposed to use b-modulation for the continuous spectrum, which was later considered also in [73, 90, 91]. In [89, 90], the main advantage of b-modulation is that the optical signal—the signal after the BNFT—is time limited, i.e., it does not show the typical tail of NIS signal, thus avoiding truncation errors and/or reducing the spectral efficiency [89, 90].

As far as it concerns b-modulation for the continuous spectrum, it has been discussed in the framework of continuous-only spectrum modulation[5]. Indeed,

[5]On the one hand, the possibility of using b-modulation with discrete spectrum (and, thus,

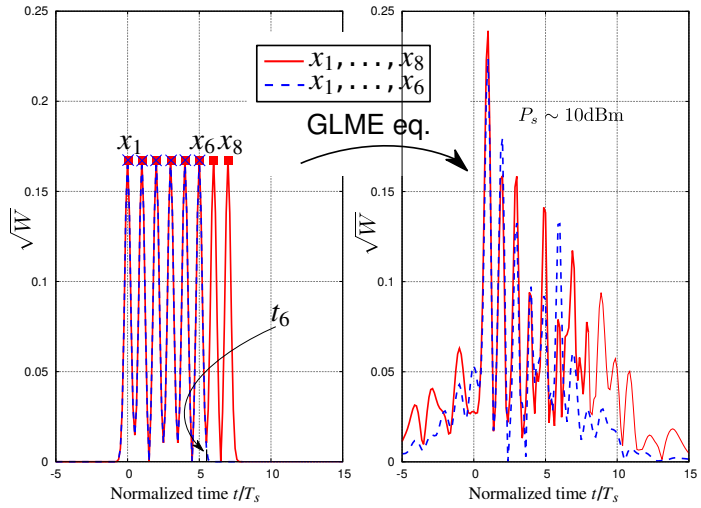


Figure 4.30: The Corollary 13 given by the NFT causality property does not hold for exponential mapping. A train of Gaussian pulses, modulated by QPSK symbols, and almost ISI-free, is shown before (on the left) and after (on the right) the BNFT. The red signal is generated by 8 symbols, while for the blue one only the first 6 are taken into account. The two optical signals are not superimposed (baudrate $R_s = 50\,\text{GBd}$, optical power $P_s = 10\,\text{dBm}$).

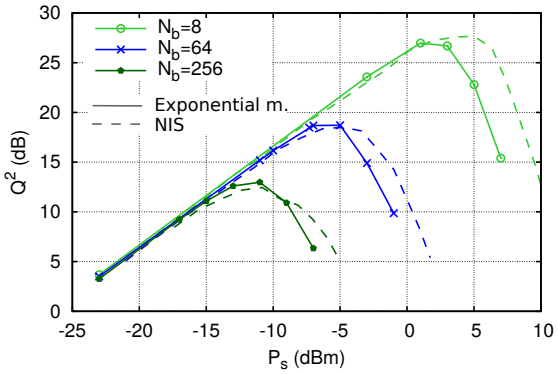


Figure 4.31: Q-factor vs optical launch power for standard NIS (dashed lines) and exponential mapping (solid lines and marks). Same scenario as in Fig. 4.9, i.e., QPSK symbols, $\beta_2 = -20.39\,\text{ps}^2/\text{km}$, $N_z = 800$, $L = 2000\,\text{km}$, and $R_s = 50\,\text{GBd}$.

in this case both $a(\lambda)$ and $\log(a(\lambda))$ are analytic and, therefore, the phase of $a(\lambda)$ can be obtained from its modulus with [12, 89]

$$\mathrm{Arg}(a(\lambda)) = \mathcal{H}(\log(|a(\lambda)|)) \tag{4.39}$$

where $\mathcal{H}$ is the Hilbert transform. Therefore, one can obtain the optical signal with the following steps:

1. encode the information on $b(\lambda)$, e.g., as in NIS with $b(\lambda) = -S(-\lambda/\pi)$ from Eq. (4.2);
2. obtain $|a(\lambda)| = (1 - \sigma|b(\lambda)|^2)^{1/2}$, which follows from Proposition 7;
3. obtain the phase of $a(\lambda)$, and thus $a(\lambda)$ itself using Eq. (4.39);
4. define $\rho(\lambda) = b(\lambda)/a(\lambda)$;
5. perform a BNFT to obtain the time domain signal.

At the RX, one only need to compute the scattering data $b(\lambda)$. Unfortunately, the requirement (see Proposition 9)

$$\sigma|b(\lambda)|^2 \leq 1, \tag{4.40}$$

imposes significant limitations in the step 1, and, thus, on the power of the time domain signal [60]. For example, if one considers NIS scheme in the same scenario as Fig. 4.9, one can only achieve the power level shown in Fig. 4.32. Importantly, the energy barrier depends on the modulation of the initial signal $s(t)$, and can be improved by proper design [73, 89].

Nevertheless, an important advantage of b-modulation is that the RX avoids to compute the nonlinear spectrum through division by $a(\lambda)$ (or $a'(\lambda)$ for the discrete spectrum). Consequently, (i) only the noise on $b(\lambda)$ affects system performance, rather than summing up with that on $a(\lambda)$ as in [72], and (ii) the RX avoids division by a number with small modulus. Indeed, despite the fact that $0 < |a(\lambda)| \leq 1$, when power increases and in presence of noise, $|a(\lambda)|$ might tend to zero at some frequencies, causing some peaks on $\rho(\lambda)$. This effect, that we conjectured being related both to numerical issues and the rising of discrete components, caused the peaks in Fig. 4.10(a) [75, 78].

To overcome power limitations caused by (4.40), one might combine b-modulation with exponential mapping, as suggested in [91], and described in the following.

$\sigma = 1$) is feasible obtaining $a(\lambda)$ as [92, 93]

$$a(\lambda) = \sqrt{1 - |b(\lambda)|^2} \exp\left\{ j\mathcal{H}\left(\log\left(\sqrt{1 - |b(\lambda)|^2} \right) \right) \right\} \prod_{k=1}^{N_{\mathrm{DS}}} \frac{\lambda - \lambda_k}{\lambda - \lambda_k^*} \tag{4.38}$$

where $\{\lambda_k\}_{k=1}^{N_{\mathrm{DS}}}$ are the discrete eigenvalues. If $N_{\mathrm{DS}} = 0$, this is equivalent to the continuous-only spectrum modulation. On the other hand, the authors of [73] manifested some doubts about the possibility of merging b-modulation on the continuous spectrum with discrete spectrum.

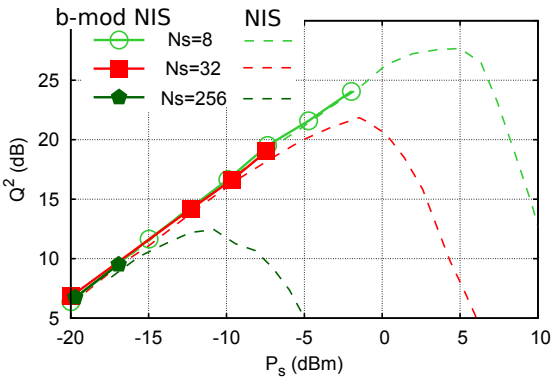


Figure 4.32: Q-factor vs optical launch power for standard NIS (dashed lines) and b-modulation with NIS mapping (solid lines and marks). Same scenario as in Fig. 4.9. QPSK symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 800$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$. Higher powers could not be obtained in this scenario due to (4.40).

1. Encode the information on $s(t)$, and obtain $S(f)$ as in NIS.
2. Map $S(f)$ on $b(\lambda)$with
$$b(\lambda) = \sqrt{\sigma(1 - e^{-\sigma|S(\lambda/\pi)|^2})}e^{j\mathrm{Arg}(S(\lambda/\pi))} \tag{4.41}$$
such that $\sigma|b(\lambda)| \leq 1$ for any $\lambda \in \mathbb{R}$. Notice that the sign of σ changed with respect to Eq. (4.34), this is done to fulfill (4.40).
3. Obtain $a(\lambda)$ from $b(\lambda)$, as previously. Define the nonlinear spectrum $\rho(\lambda) = b(\lambda)/a(\lambda)$.
4. Perform a BNFT to obtain the time domain signal $q(t)$.

At the RX, one needs to compute $b(\lambda)$ and invert Eq. (4.41) through $S(f) = \sqrt{-\sigma\log(1-\sigma|b(\pi f)|^2)}e^{j\mathrm{Arg}(b(\pi f))}$. Also in this case, the energy is maintained between the QAM signal and the optical signal (since $1 + \sigma|\rho|^2 = 1/|a|^2 = (1-\sigma|b|^2)^{-1}$).

Unfortunately, the time domain signal obtained with b-modulation and exponential mapping does not have the same time duration of the initial signal $s(t)$ and, therefore, an important advantage of b-modulation is lost. However, it worth mentioning that in many cases (e.g., the scenario considered in this Section) the loss in spectral efficiency due to the NFT boundary conditions is much smaller than that induced by dispersion. Conversely, the advantages deriving from decoding information from $b(\lambda)$ only is maintained here. The performance of b-modulation in combination with exponential mapping is shown in Fig. 4.33, where it is compared with the performance of NIS in the same scenario as in Fig. 4.9. The figure shows that b-modulation combined with exponential

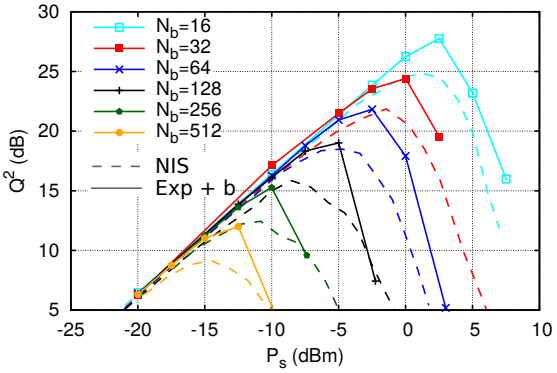


Figure 4.33: Q-factor vs optical launch power for standard NIS with dashed lines and b-modulation with exponential mapping with increased accuracy with solid lines ($N_F = N_B = 8$). Same scenario as in Fig. 4.9, i.e., QPSK symbols, $\beta_2 = -20.39\,\mathrm{ps}^2/\mathrm{km}$, $N_z = 800$, $L = 2000\,\mathrm{km}$, and $R_s = 50\,\mathrm{GBd}$.

mapping achieves $\sim 3\,\mathrm{dB}$ improvement. Importantly, the numerical accuracy required by this mapping is higher (the performance worsens when considering $N_F = N_B = 4$ as for NIS).

Recently, a method to use b-modulation with compact signals and negligible power limitation has been proposed in [73].

4.7 Conclusion and outlook

This Chapter dealt with single-polarization NFDM schemes that encode the information on the continuous part of the nonlinear spectrum—using the NFT for the localized NLSE. This Chapter analyzed NFDM schemes in a simple scenario—single-user, single-polarization, lossless channel—to provide a contribution from a theoretical point of view. The work has been driven by the future, rather than imminent, goal of optimizing NFDM to possibly outperform conventional systems.

After a general introduction about NFDM, this Chapter focused on NIS scheme and highlighted some issues related to the unusual—for conventional linear systems—signal–noise interaction in NIS. Some solutions—namely, windowing and precompensation—have been proposed, which can partly mitigate NIS weaknesses. Next, the Chapter focused on detection strategies, with a twofold aim: (i) show that current NIS, and more in general NFDM, is not optimized and significant gain can be obtained by designing tailored transmission paradigms, (ii) propose, analyze and discuss novel detection strategies for NIS, which provided significant performance improvements. Finally, we briefly considered other kind of NFDM paradigm recently suggested, and compared those with NIS. We showed that b-modulation with exponential mapping performs better than NIS.

Despite the contribution of this Chapter, and many other works from different research groups, NFDM schemes have not yet been fully understood and optimized. To understand the statistics of noise in the nonlinear spectrum, define optimal detection strategies as well as encoding techniques, are just some of the open research topics in the field. Furthermore, it is worth mentioning that (i) fast NFT algorithms, and (ii) nonlinear add-drop multiplexer (NADM) [11]—rather than standard reconfigurable optical add-drop multiplexer (ROADM)—are essential for the advent of NFDM transmission scheme in the real word.

Chapter 5

Dual-polarization nonlinear frequency-division multiplexing systems

Until recently, NFT-based transmission schemes have been mostly considered in the single-polarization case (see Chapter 4) and, hence, based on the NFT processing associated with the NLSE channel ($\mathrm{NFT}_{\mathrm{NLS}}$ in the following of this Chapter). However, the SMF supports two orthogonal propagation modes and high-efficiency transmission methods typically use both polarization components for modulation. Under some realistic conditions, the averaged dynamics of two orthogonal modes in randomly-birefringent fibers is governed by the integrable version of the ME Eq. (2.8) [31] (see Chapter 2), whose NFT form ($\mathrm{NFT}_{\mathrm{M}}$ in the following) has been known since the original paper by Manakov [33] (see Chapter 3). The possibility to double the transmission rate of NFT-based systems by employing both polarization components had remained almost unexplored until 2017 aside, perhaps, from just one earlier work [71]. At the same time, the need of incorporating both polarization components into NFT-based systems is apparent, such that, more recently, joint polarization and nonlinear frequency-division multiplexing (PNFDM) schemes have been gradually getting more attention [26,40,49,68,70,71,87,88,91,94–97]. This chapter deals with the polarization-division multiplexing NIS (PDM-NIS), the dual-polarization analog of the NIS scheme that was initially proposed for the NLSE in [19, 50] and detailed in Chapter 4.

In this Chapter, the BER is estimated through the EVM. The $\mathrm{FNFT}_{\mathrm{M}}$ and $\mathrm{BNFT}_{\mathrm{M}}$ operations are implemented by using the numerical methods presented in Section 3.4, considering the case $M = 2$ for the ME. Unless otherwise stated, an oversampling factor of $N_F = N_B = 4$ samples per symbols is used.

The Chapter is organized as follows. Section 5.1 describes the PDM-NIS system: the performance is studied and compared with that of single-polarization NIS systems reported in Section 4.2. Next, Section 5.2 investigates a simpli-

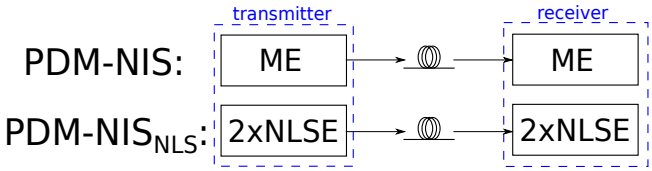


Figure 5.1: Details of NFT processing in PDM-NIS (above) and PDM-NIS$_{\mathrm{NLS}}$ (below).

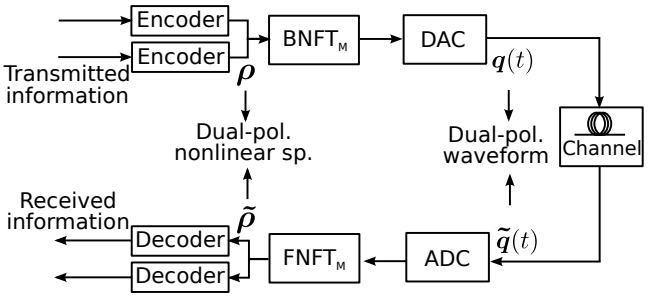


Figure 5.2: Basic PDM-NIS scheme.

fied approach to PDM-NIS—the PDM-NIS$_{\mathrm{NLS}}$ scheme—which is based on the scalar NFT$_{\mathrm{NLS}}$ processing. The difference between PDM-NIS and PDM-NIS$_{\mathrm{NLS}}$ is highlighted in Fig. 5.1. Finally, the Conclusion section draws the conclusion and outlines possible future works. This Chapter reports some of the results presented in [40, 95, 96].

5.1 Dual-polarization nonlinear frequency-division multiplexing setup and performance

The system setup is sketched in Fig. 5.2, and is the natural dual-polarization extension of the NIS scheme considered in Section 4.2 [75]. At the TX, information is mapped on two QPSK signals $s_i(t)$, $i = 1, 2$, with a shaping pulse having a root raised cosine (RRC) FT with roll-off $\beta = 0.2$, and symbol rate $R_s = 50\,\mathrm{GBd}$. After normalization, the FT of each $s_i(t)$, $S_i(f)$, is mapped to the nonlinear spectrum (3.26) according to $\rho_i(\lambda) = -S_i(-\lambda/\pi)$, for $i = 1, 2$. The dual-polarization optical signal $\mathbf{q}(t)$ is obtained performing an BNFT$_{\mathrm{M}}$ of the dual-polarization nonlinear spectrum $\boldsymbol{\rho}(t) = (\rho_1(\lambda), \rho_2(\lambda))$. Next, the analog signal is obtained with a DAC and sent into the channel. The channel is the same considered in Section 4.2. A preliminary study about the impact of PMD on the NFT-based transmission showed that it can be compensated with a very small performance degradation [49], such that we neglect the impact of PMD in

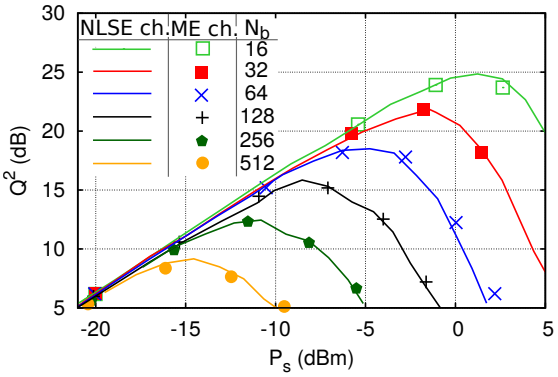


Figure 5.3: Performance (Q^2-factor) of single-polarization modulation over the ME channel model (symbols only), where both polarization components are corrupted by noise, compared to that over the NLSE channel model (solid lines) (same as in Section 4.2).

the current work. At the end of the channel, the ADC recovers the samples of received signal, from which the received nonlinear spectrum is retrieved through the $\mathrm{FNFT_M}$ block. A noise-corrupted version of $\rho_i(\lambda)$, $i = 1, 2$, is obtained from the received signal and multiplied by $e^{4j\lambda^2\mathcal{L}}$ to remove the deterministic propagation effects (for the sake of simplicity, we do not use precompensation as in Section 4.4, but it can be done straightforwardly). Finally, matched filtering and sampling are used to recover the transmitted information symbols. Both the DAC and ADC have bandwidth $B = 100\,\mathrm{GHz}$. As in Section 4.2, $N_z = 800$ is considered to account for the overall memory due to linear dispersion. It is important to remark that, while the operations concerning symbol mapping (detection) on (from) the nonlinear spectrum are performed independently on the two polarizations (and are the same considered in Section 4.2), $\mathrm{FNFT_M}$ and $\mathrm{BNFT_M}$ are performed jointly (and the result depends on both polarizations) to ensure the integrability of the channel (2.8).

Most of the works dealing with the NFT-based transmission schemes consider the NLSE (single-polarization) channel model for simulations. However, in practical transmission systems, in-line amplifiers generate noise on both polarizations, thus making the second polarization $q_2(t)$ always non null. Therefore, the two polarizations can interact with each other due to the nonlinear coupling term present in the ME. To investigate the possible impact of this coupling, Fig. 5.3 compares the performance obtained with the single-polarization NIS scheme assuming the NLSE as a channel model, with that obtained with the same scheme but assuming the full ME as a channel model. Thus, only one polarization of the ME is modulated, while the other grows during propagation due to in-line noise, and is eventually discarded at the RX. The figure shows that the systems performance does not change noticeably, meaning that the noise in the second polarization does not affect the NIS performance. Note that

it might not be the case for other parameter ranges, transmission schemes, or detection strategies.

In Fig. 5.4(a) we show with solid lines the PDM-NIS performance as a function of the launch power P_s in Eq. (2.19) for different burst lengths N_b. For the sake of comparison, the dashed lines in the same figure show the results obtained in the same system when we modulate only one polarization and set the other one to zero. We remark that, while the same colors correspond to the same burst lengths N_b, the number of information symbols is doubled when considering the PDM-NIS compared to the single-polarization NIS. The figure shows that the PDM-NIS performance is slightly worse than that obtained for single-polarization NIS. This difference increases up to about 1 dB for longer bursts. We conjecture that this degradation is due to the doubled energy of the received signal, which might affect the strength of the perturbation caused by noise on the nonlinear spectrum. Indeed, some theoretical studies [15, 64] indicate that, when considering the NLSE model, the intensity of the noise affecting the nonlinear spectrum increases with the spectrum itself. However, to the best of our knowledge, similar studies are not available for the ME. Finally, notice that Fig. 5.4(a) reports on the xaxis the average input optical power *per symbol* defined through Eqs. (2.18)-(2.19), meaning that the total average input optical power in the dual polarization system would be doubled with respect to the single polarization case, as shown for example in [49].

Importantly, the decay of PDM-NIS (and NIS) performance with the burst length is caused by noise and not by numerical inaccuracies, as demonstrated in the following. In fact, Fig. 5.4(b) compares the performance of PDM-NIS in a noisy and ideal noise-free scenario, and with actual ($N_F = N_B = 4$) and increased ($N_F = N_B = 8$) numerical accuracy for the BNFT and FNFT computation. The decay of the noise-free performance at higher power is a typical behavior of NFT-based schemes, and is due to the fact that, at higher powers, the system is more sensitive to numerical inaccuracies. Consequently, a higher numerical accuracy provides a better performance in the noise-free scenario. On the other hand, in the noisy scenario, the impact of noise is much stronger than that of numerical inaccuracy (as testified by the significant performance decrease compared to the noise-free scenario), such that PDM-NIS achieves the same performance with standard or increased numerical accuracy. Similar conclusions were drawn in Chapter 4.

Furthermore, to understand the impact of numerical errors in the NFT operations, and whether the FNFT or the BNFT has more impact, we studied the error incurred when performing these operations. In particular, we considered the samples of the nonlinear spectrum $\boldsymbol{\rho}(\lambda)$ obtained for PDM-NIS (same scenario considered in Figs. 5.4(a) and 5.4(b) with $N_b = 32$), and we applied one BNFT and one FNFT to obtain the samples of

$$\tilde{\boldsymbol{\rho}}(\lambda) = \mathrm{FNFT}(\mathrm{BNFT}(\boldsymbol{\rho}(\lambda)). \tag{5.1}$$

Figure 5.5(a) shows the normalized mean square error (NMSE) on the nonlinear

spectrum defined as

$$\text{NMSE} = \left(\sum_{m=1,2} \sum_{k=1}^{N_{\text{sa}}} |\rho_{k,m} - \tilde{\rho}_{k,m}|^2 \right) \left(\sum_{m=1,2} \sum_{k=1}^{N_{\text{sa}}} |\rho_{k,m}|^2 \right)^{-1} \tag{5.2}$$

where N_{sa} is the number of samples for the nonlinear frequency λ, and $\rho_{k,m}$ and $\tilde{\rho}_{k,m}$ are the k-th samples of $\rho_m(\lambda)$ and $\tilde{\rho}_m(\lambda)$, respectively, with $m = 1, 2$ indicating the two polarizations. The figure shows the NMSE for different oversampling factors for the FNFT and the BNFT—N_F and N_B, respectively. The blue curve represents the error obtained with actual accuracy ($N_F = N_B = 4$), while the red one ($N_F = N_B = 16$) can be taken as "reference"; the impact on error of the two NFT operation is shown by decreasing one at a time N_F and N_B from the reference one. It turns out that increasing the oversampling of the FNFT (i.e., N_F) improves the accuracy mostly in the higher power region, while increasing the oversampling for the BNFT (i.e., N_B) mostly reduces the error in the lower power region.

Furthermore, the numerical algorithm used for the FNFT is not exact, but uses the approximated approach discussed in Section 3.4.1. We investigate the impact of this choice in Figs. 5.5(b). The figure shows that the performance of the PDM NIS scheme considered in Figs. 5.4(a)-(b) are unaltered (also in the noise-free scenario) when considering the exact method, revealing that the FNFT algorithm is not affecting the system performance. Secondly, to understand the difference between the two algorithms—i.e., using the approximation to the transfer matrix (3.82) or the exact transfer matrix (3.88)—Figs. 5.6(a)-(b) plot the NMSE (5.2) as a function of the oversampling factor for the FNFT, and considering an ideal BNFT (i.e., with high enough accuracy). The figure shows that for lower powers (Fig. 5.6(a)), the two algorithms perform the same. Conversely, for higher powers (Fig. 5.6(b)), the exact algorithm is more accurate, in particular for higher oversampling factors. This means that depending on the circumstances, the exact algorithms may have better converges properties. However, in the scenario considered in this Chapter, the non-exact algorithm can be used without any penalty. Moreover, when considering the speed of convergence of the two algorithms, the computational cost should also be taken into account, see Section 3.4.1.

5.2 Reduced complexity system with nonlinear Schrödinger equation processing

The ME (3.14) describes the propagation of a normalized dual-polarization optical signal in the fiber channel, accounting for the interaction between the two polarizations induced by the nonlinear term (see Section 2.2.2). Accordingly, the PDM-NIS encodes and decodes information on the nonlinear spectrum using the NFT_{M} associated with the ME, as in Fig. 5.2, avoiding nonlinear interference. The ME does not entail any exchange of energy between the two signal

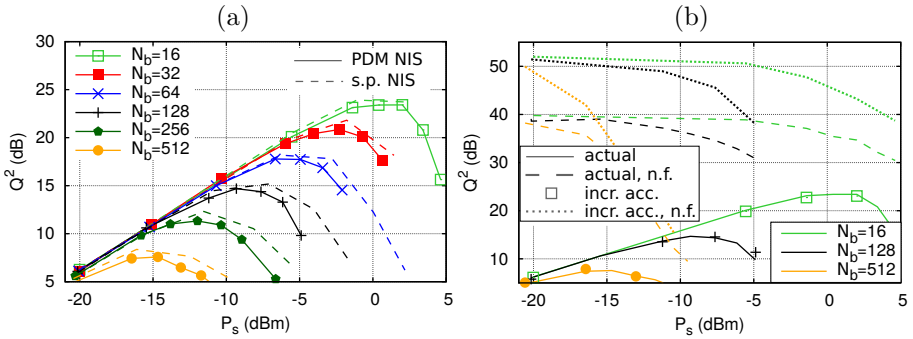


Figure 5.4: Performance for different burst lengths with same color: (a) PDM-NIS performance compared with single-polarization NIS; and (b) PDM-NIS in the noisy and noise-free (n.f.) scenarios, with actual ($N_F = N_B = 4$ samples per symbol) and increased ($N_F = N_B = 8$ samples per symbol) accuracy for NFTs.

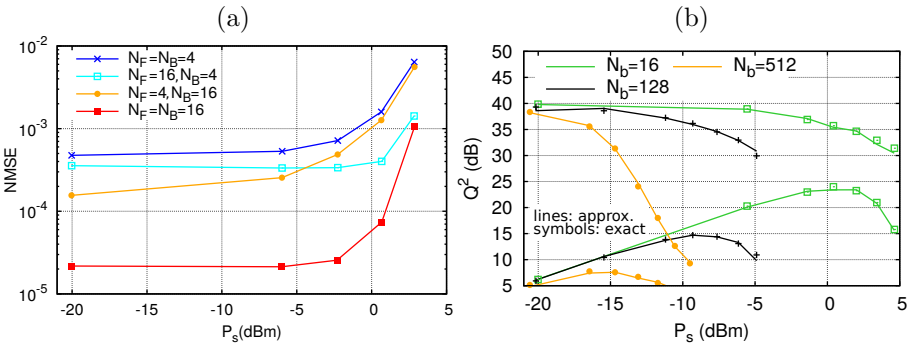


Figure 5.5: (a) NMSE on the nonlinear spectrum after BNFT and FNFT as a function of the optical power, for different oversampling factors. (b) system performance with channel and in B2B with the two FNFT methods presented in Section 3.4.1; same scenario as Fig. 5.4.

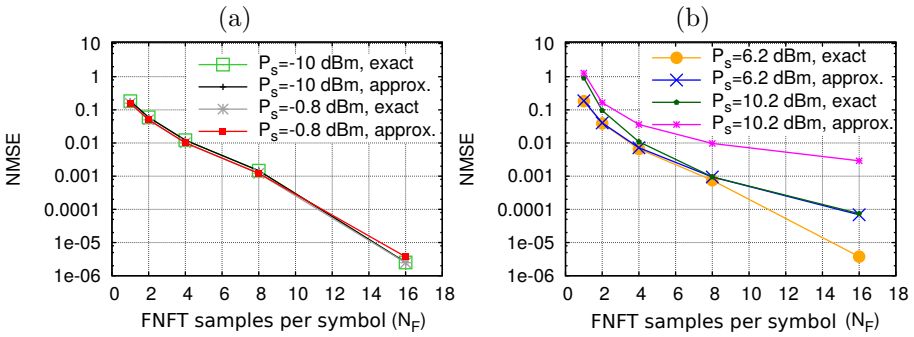


Figure 5.6: NMSE on the nonlinear spectrum after BNFT and FNFT as a function of the oversampling factor for the FNFT N_F for different power level (a) $P_s = -10\,\text{dBm}$, $-0.8\,\text{dBm}$, and (b) $P_s = 6.2\,\text{dBm}$, $10.2\,\text{dBm}$. The oversampling factor for the BNFT is $N_B = 16$, the number of symbols is $N_s = 8$.

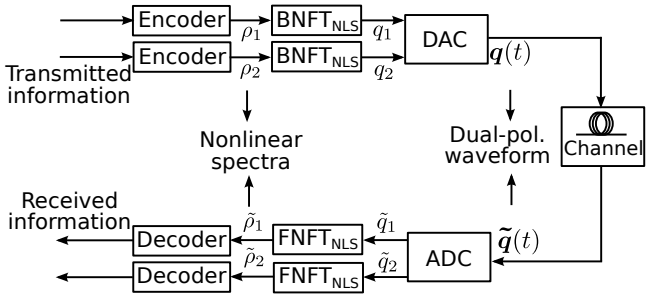


Figure 5.7: PDM-NIS$_{\text{NLS}}$ scheme.

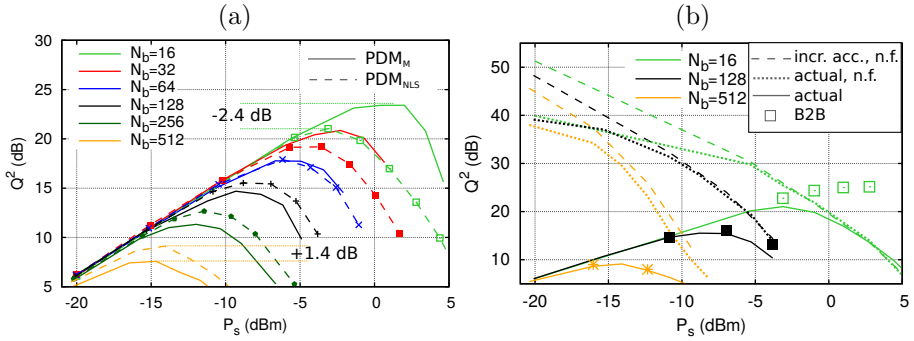


Figure 5.8: Performance vs power per symbol for different burst lengths with same color: (a) PDM-NIS (solid lines) compared with reduced complexity PDM-NIS$_{\text{NLS}}$ (dashed lines); (b) PDM-NIS$_{\text{NLS}}$ compared with PDM-NIS$_{\text{NLS}}$ noise-free (n.f.) with actual ($N_F = N_B = 4$ samples per symbol) and increased ($N_F = N_B = 8$ samples per symbol) accuracy, and with B2B performance.

polarizations, which suggests that modeling their propagation by two independent NLSEs might provide a reasonable approximation. In this case, the NIS transmission scheme could be implemented independently on each polarization according to the PDM-NIS$_{\mathrm{NLS}}$ scheme shown in Fig. 5.7. This approximated approach neglects the interaction between the two polarizations during propagation, giving rise to some nonlinear interference. At the same time, using two NFT$_{\mathrm{NLS}}$ instead of a single NFT$_{\mathrm{M}}$ reduces the overall processing complexity, as will be clear later in this section. It is therefore interesting to see what is the impact of the introduced simplification on the performance of the NIS system.

To address this problem, we compare the achievable performance of PDM-NIS with PDM-NIS$_{\mathrm{NLS}}$ depicted, respectively, in Figs. 5.2 and 5.7. Simulation results are shown in Fig. 5.8(a) for different burst lengths. At lower powers, the performance of PDM-NIS$_{\mathrm{NLS}}$ and PDM-NIS is the same. Indeed, in the linear regime, the nonlinear term in the ME (3.3), which accounts for polarization mixing, tends to zero. Consequently, the two transmission schemes are equivalent. At higher powers, the two schemes perform differently. For shorter bursts (i.e., $N_b = 16, 32$) PDM-NIS$_{\mathrm{NLS}}$ performs worse, as expected, due to the mismatch between the transmission scheme (designed for the NLSE) and the actual channel (modeled by the ME). On the other hand, increasing the burst length, the performance difference decreases and, for long burst (i.e., $N_b = 128, 256, 512$), PDM-NIS$_{\mathrm{NLS}}$ performs even slightly better than PDM-NIS. We conjecture that this unexpected behavior has the same physical origin as the performance degradation of PDM-NIS compared to single-polarization NIS observed in Fig. 5.4(a). Indeed, in PDM-NIS$_{\mathrm{NLS}}$, detection is made by separately considering the NFT$_{\mathrm{NLS}}$ spectrum of each polarization, whose energy is only one half that of the total signal. Therefore, recalling that the intensity of the perturbation of the nonlinear spectrum caused by noise depends on the signal energy, we expect the NFT$_{\mathrm{NLS}}$ spectrum of each polarization to be less affected by noise than the NFT$_{\mathrm{M}}$ spectrum of the whole signal. This effect is more evident for higher signal energies, i.e., for longer bursts, when it becomes stronger than the mismatch between the transmission scheme and the channel. This outcome shows that, in the region where this effect is evident, signal noise interaction in the joint processing strongly affects performance and hides the benefit of considering the ME to include polarization interaction. Finally, it is worth noting that for longer bursts, PDM-NIS$_{\mathrm{NLS}}$ performs similarly to the single-polarization NIS, cf. Fig. 5.4(a).

Figure 5.8(b), which shows the performance of PDM-NIS$_{\mathrm{NLS}}$ (i) with dotted line, (ii) in the ideal noise-free scenario with solid line, (iii) in the ideal noise-free scenario and with increased accuracy for the NFTs with dashed lines, and (iv) in the B2B configuration with symbols only, supports our conjecture, as explained in the following. Firstly, Fig. 5.8(b) shows that at higher powers the PDM-NIS$_{\mathrm{NLS}}$ performance coincides with the noise-free performance, indicating that the performance decay does not originate from noise. Secondly, the performance of PDM-NIS noise-free, which is shown in Fig. 5.4(b), equals that of PDM-NIS$_{\mathrm{NLS}}$ at lower powers, but PDM-NIS performs better at higher

powers, indicating that the system does not account for the polarization mixing occurring at high powers. Thirdly, when increasing accuracy, the performance of PDM-NIS$_{\text{NLS}}$ noise-free increases at lower powers, where the polarization mixing is negligible, but does not improve at higher powers. Moreover, the performance improves for $N_b = 16, 128$ in the B2B configuration, i.e, without channel, but with an equivalent noise. The latter two facts confirm that the performance degradation occurs due to the polarizations' interaction.

The impact of the mismatch [96] in PDM-NIS$_{\text{NLS}}$ is shown in Fig. 5.9(a), which compares the performance with the actual fiber and in B2B configuration (but with same total accumulated noise), as a function of P_s. For smaller N_b, the B2B performance (when there is no propagation and, hence, no mismatch) is significantly higher than after 2000 km (when the mismatch becomes relevant). Conversely, for larger N_b, the noise impact on the system is stronger than the mismatch (recall that the impact of noise on the nonlinear spectrum increases with N_b, as demonstrated in the single-polarization case [15, 64, 75]), such that the latter becomes negligible and the same performance is achieved in B2B and after 2000 km. In this case, PDM-NIS$_{\text{NLS}}$ performs better than PDM-NIS, as shown in Fig. 5.8(a). To further confirm this hypothesis, Fig. 5.9(b) reports the optimal performance of PDM-NIS and PDM-NIS$_{\text{NLS}}$ as a function of the propagation length L (at the optimal launch power for each curve), and for the same overall accumulated noise. While this scenario is not realistic—the noise changes with transmission length—it helps to highlight the impact of the mismatch occurring during propagation. Firstly, the figure shows that the performance of PDM-NIS remains unchanged, similarly to NIS (see Section 4.2). Secondly, the figure confirms the presence of a threshold between the impact of noise (constant in L in this particular scenario) and the impact of the deterministic mismatch in PDM-NIS$_{\text{NLS}}$ (which increases with L): when the first is more significant, PDM-NIS$_{\text{NLS}}$ has to be preferred, while PDM-NIS performs better in the other case. Finally, Fig. 5.10 compares PDM-NIS and PDM-NIS$_{\text{NLS}}$ as a function of the propagation length (at the optimal launch power for PDM-NIS$_{\text{NLS}}$) and shows that their reciprocal behavior remains unchanged for different lengths.

Figure 5.11(a) compares for $N_b = 32$ the performance of PDM-NIS and PDM-NIS$_{\text{NLS}}$ with those of a conventional system—i.e., without any NFT—employing EDC or DBP with 1 or 10 step per span. The figure shows that PDM-NIS and PDM-NIS$_{\text{NLS}}$ both outperform EDC, while DBP with 1 step per span is comparable with PDM-NIS; DBP with 10 step per span outperforms the other schemes. This result is in accordance with that obtained for single-polarization in [75], and we expect to obtain the same behavior shown in [75] for different values of N_b. Also, a comparison with conventional systems for dual-polarization systems has been shown in [49], which reports results more favorable for NFT based schemes. However, we mention that the NFDM schemes are expected to provide the best improvements with respect to conventional systems when the multi-channel transmission in the network scenario with ROADMs is considered, while here we consider a single channel (a point-to-point transmission). Also, while the final goal of NFDM systems is to outperform conventional systems overcoming the limitations imposed

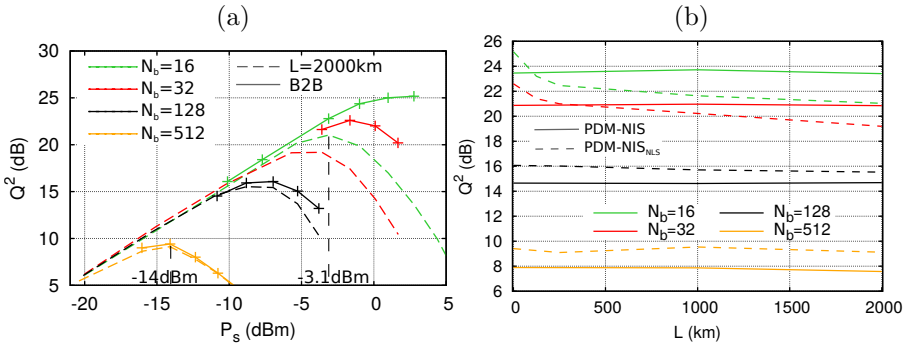


Figure 5.9: (a) PMD-NIS$_{\text{NLS}}$ performance vs power P_s with channel and in B2B configuration; (b) Optimal Q-factor vs transmission length for fixed power P_s with the same overall noise for PDM-NIS and PDM-NIS$_{\text{NLS}}$.

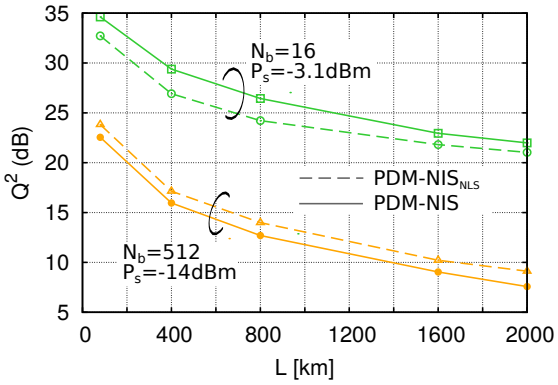


Figure 5.10: PDM-NIS and PMD-NIS$_{\text{NLS}}$ performance vs propagation length L for fixed optimal powers.

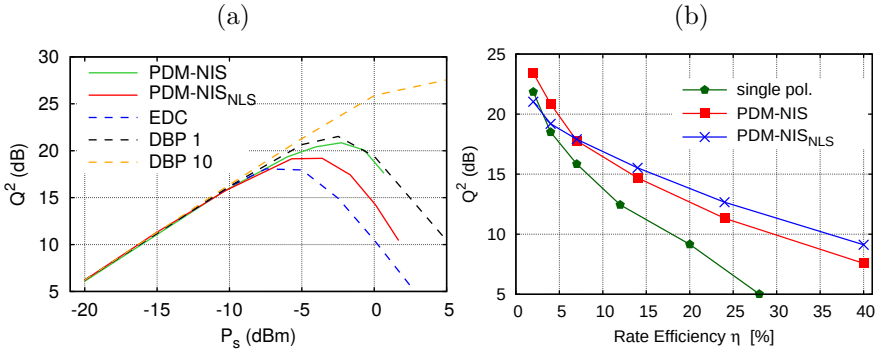


Figure 5.11: (a) Performance vs power per symbol for $N_b = 32$ for PDM-NIS, PDM-NIS$_{\text{NLS}}$, and conventional systems (EDC and DBP with 1 and 10 step per span); (b) Optimal performance as a function of the rate efficiency.

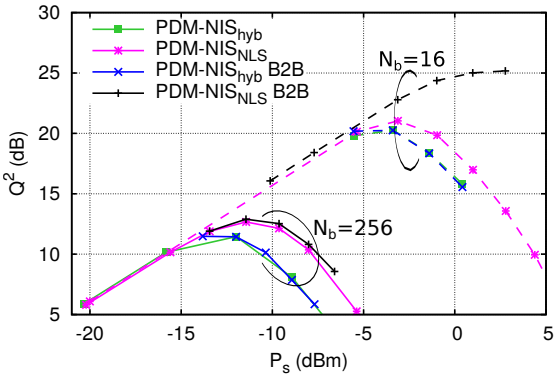


Figure 5.12: Performance vs power per symbol for $N_b = 16, 256$ for PDM-NIS$_{\mathrm{NLS}}$ and PDM-NIS$_{\mathrm{hyb}}$ with the actual channel and in B2B configuration.

by nonlinearity, here we investigate dual-polarization NIS schemes to provide a tool that, once optimized, might compete with conventional systems. An interested reader can find more comparisons between PDM NFT-based systems and OFDM in [49] and comparisons in single-polarization in [18, 50, 75] and in Section 4.2 [75].

Figure 5.11(b) compares the performance of single-polarization NIS, PDM-NIS, and PDM-NIS$_{\mathrm{NLS}}$ as a function of the rate efficiency (2.17), which accounts for the loss in spectral efficiency due to the insertion of guard times, and the overall number of information symbols sent [75]. Firstly, Fig. 5.11(b) shows that, thanks to the use of both polarizations, PDM-NIS performs better than single-polarization NIS, doubling the rate efficiency with only a small performance degradation. Secondly, for lower rate efficiencies, PDM-NIS$_{\mathrm{NLS}}$ performs worse than both dual and single-polarization NIS, as a result of neglecting polarizations' interaction. On the other hand, at higher rate efficiencies, PDM-NIS$_{\mathrm{NLS}}$ performs slightly better (around 1 dB) even than PDM-NIS, thanks to the lower impact of noise on the NFT$_{\mathrm{NLS}}$ spectrum. In this work precompensation is not deployed, but can be used to halve the number of guard symbols N_z and, thus, increase the spectral efficiency [79, 85]. This, however, would not change the overall behavior of Fig. 5.11(b).

Following the reasoning considered in this section, one may be tempted to consider an hybrid scheme (PDM-NIS$_{\mathrm{hyb}}$), which (i) modulates the information according to the NFT$_{\mathrm{M}}$, i.e., in agreement with the channel model, and (ii) retrieves the information using two NFT$_{\mathrm{NLS}}$, to reduce the impact of noise on the nonlinear spectrum. Figure 5.12 compares the performance obtained with PDM-NIS$_{\mathrm{hyb}}$ and PDM-NIS$_{\mathrm{NLS}}$, considering also the B2B scenario. The figure shows that PDM-NIS$_{\mathrm{hyb}}$ performs worse than PDM-NIS$_{\mathrm{NLS}}$, and its performance is superimposed with that in the B2B configuration, differently from PDM-NIS$_{\mathrm{NLS}}$. This suggests that, while in PDM-NIS$_{\mathrm{NLS}}$ the TX and the RX agree with each others and errors occur because of the presence of the channel,

the hybrid scheme already introduces a discrepancy in the B2B configuration, since the TX and the RX are not matched. As a consequence, this transmission scheme (which does not provide a significant complexity reduction) is not comparable with PDM-NIS$_{\mathrm{NLS}}$ in terms of performance.

We note that, while the NFT$_{\mathrm{M}}$ theory required for double-polarization NFT-based communication systems can be deemed a straightforward extension of the NFT$_{\mathrm{NLS}}$ theory, it can bring about some difficulties in terms of developing fast and accurate numerical algorithms for NFT$_{\mathrm{M}}$ computation, in particular taking into account that the research for fast numerical NFT$_{\mathrm{NLS}}$ is still in progress (see [7] and references therein). Indeed, the computational complexity depends on the algorithms deployed and further work is required in this direction. However, we expect the computational complexity of PDM-NIS$_{\mathrm{NLS}}$ to be typically lower than that of PDM-NIS because of an extra dimension entering the operations involved in the latter. These aspects might become even more relevant when increasing the number of dimensions, e.g., by extending the PDM-NIS concept and the complexity reduction approach based on PDM-NIS$_{\mathrm{NLS}}$ to SDM systems in multi-core or multi-mode fibers. Indeed, considering the general case with $M \geq 2$, the RX should solve M 2×2 or one $(M+1) \times (M+1)$ Z-S eigenvalue problem. Let $\mathcal{C}(M, N_{\mathrm{sa}})$ denote the computational cost for solving the Z-S eigenvalue problem of dimension $M+1$ (i.e., corresponding to the M-dimensional VNLSE) its computational cost, $N_{\mathrm{sa}} \geq 1$ being the number of samples for the time axis. With this notation, the reduced complexity RX would be less computationally complex if and only if

$$M\mathcal{C}(1, N_{\mathrm{sa}}) < \mathcal{C}(M, N_{\mathrm{sa}}). \tag{5.3}$$

This equation is true for the exact algorithm presented in Section 3.4.1 since, in this case, $\mathcal{C}(M, N_{\mathrm{sa}}) = N_{\mathrm{sa}}(M+1)^2$, as follows from Eq. (3.96). Moreover, we expect Eq. (5.3) to hold also with faster algorithms as the discretized time domain signal has $N_{\mathrm{sa}}M$ samples and a sufficient condition for Eq. (5.3) to hold is that $\mathcal{C}(M, N_{\mathrm{sa}})$ depends more than linearly on M. Conversely, Eq. (5.3) does not hold for the approximated algorithm in Section 3.4.1, as follows from Eq. (3.95); nevertheless, in this case the method requires further investigations.

However, we recall that the reduced-complexity system performs better only in some specific scenarios (when the perturbation of the nonlinear spectrum due to noise dominates the performance) and for the considered detection strategy. In fact, we expect that when dealing with improved detection strategies which can avoid the aforementioned detrimental perturbation of the nonlinear spectrum [54, 76], a joint processing of all the system modes (polarizations) by the NFT$_{\mathrm{M}}$ might be required to obtain the optimal performance.

As an end note, we would like to remark an important difference between PDM-NIS and PDM-NIS$_{\mathrm{NLS}}$, which regards the (slowly varying in time) polarization rotation induced on a signal during propagation, which can be modeled as multiplication for an unitary matrix R, obtaining $\mathbf{q}R^T$. As far as it concerns the PDM-NIS scheme, this rotation can be removed both in time (i.e., before the NFT) or in the nonlinear frequency domain (i.e., after the NFT), multiplying

(from the right) for $R^{T-1} = R^*$ or multiplying (from the left) for $R^{*-1} = R^T$, respectively. This follows from Eq. (3.49). On the other hand, the same can not be done for PDM-NIS$_{\mathrm{NLS}}$ since a property similar to Eq. (3.49) does not hold. Indeed,

$$\begin{bmatrix} \mathrm{FNFT}_{\mathrm{NLS}}\left(R_{11}q_1 + R_{12}q_2\right) \\ \mathrm{FNFT}_{\mathrm{NLS}}\left(R_{21}q_1 + R_{22}q_2\right) \end{bmatrix} \neq \begin{bmatrix} R_{11}^*\mathrm{FNFT}_{\mathrm{NLS}}\left(q_1\right) + R_{12}^*\mathrm{FNFT}_{\mathrm{NLS}}\left(q_2\right) \\ R_{21}^*\mathrm{FNFT}_{\mathrm{NLS}}\left(q_1\right) + R_{22}^*\mathrm{FNFT}_{\mathrm{NLS}}\left(q_2\right) \end{bmatrix}, \tag{5.4}$$

since the FNFT is not a linear operation and $|R_{n,m}| \neq 1$ for $n, m = 1, 2$. For a simple numerical counterexamples consider, for example, $R_{11} = R_{12} = R_{22} = 1/\sqrt{2}$and $R_{21} = -1/\sqrt{2}$. The lack of a similar property implies that the polarization rotation in PDM-NIS$_{\mathrm{NLS}}$ should be removed in time domain, before the NFT, and, therefore, digital techniques usually considered might not be extended straightforwardly in this case.

5.3 Conclusion and outlook

This Chapter dealt with dual-polarization NFT-based transmission schemes, exploiting NLSE and ME integrability. A polarization and nonlinear frequency-division scheme—PDM-NIS—was introduced, following its analogy with the NLSE-based NIS for one polarization in Chapter 4. We showed that the PDM-NIS achieves almost the same performance as we have for one-component NIS but doubling the number of information symbols transmitted. Moreover, we introduced the reduced-complexity PDM-NIS$_{\mathrm{NLS}}$ transmission scheme that, similarly to PDM-NIS, encodes and decodes information on the nonlinear spectrum, but using two scalar NFT$_{\mathrm{NLS}}$ rather than one NFT$_{\mathrm{M}}$. This scheme, which neglects polarization mixing occurring during the propagation, provides a complexity reduction, not only from a computational point of view (a lower number of floating point operations required), but also allow us to avoid the possible difficulties arising in the NFT$_{\mathrm{M}}$ theory and algorithms. Remarkably, despite the mismatch with the channel model, the performance of PDM-NIS$_{\mathrm{NLS}}$ is not only comparable with PDM-NIS at some parameters range, but can also provide a slight performance improvement at sufficiently high rate efficiency.

The same issues highlighted for NIS in Section 4.2 hold for the dual-polarization case, therefore, as a future work, we plan to extend the detection strategies presented in Section 4.5 to PDM-NIS (as the NFT property holds for the M-dimensional VNLSE). More in general, we believe that the problem of considering tailored detection strategies should be considered. In this case, it will be important to understand whether the reduced complexity scheme can still be used.

Moreover, as future works, it is important to understand the impact of adding discrete spectral components modulation to the schemes presented here as already done for single-polarization, e.g. in [18]. Also, the use of PDM-NIS and PDM-NIS$_{\mathrm{NLS}}$ within SDM-NIS schemes (multi-mode or multi-core fiber in the strong coupling regime [34,35]) should be investigated, having in mind that

the complexity of the NFT operations associated with VNLSE further increases. Finally, for the PNFDM schemes to be fully competitive with other nonlinearity mitigation techniques, it is essential to develop fast numerical algorithms for both the direct and inverse NFT operations for the ME.

Chapter 6

Experimental demonstration of dual-polarization NFDM with joint spectrum

This Chapter describes the first experimental demonstration of a dual-polarization NFDM transmission modulating both discrete and continuous nonlinear spectra [26]. Previously, there have been numerously impressive demonstrations of NFDM transmission in single-polarization [13, 16, 18, 72, 80, 82, 84, 98, 99], and, more recently, a few demonstrations of dual-polarization NFDM modulating either the continuous spectrum [87] or the discrete spectrum [68, 70, 88]. A dual-polarization transmission where both continuous and discrete spectra are jointly modulated using both polarizations represents the complete system where all the degrees of freedom for modulation provided by a SMF are exploited [26].

In this Chapter, we numerically characterize the joint BNFT framework described in Section 3.4.3 to encode the information on both polarizations and both spectra. By applying such BNFT, a dual-polarization joint NFDM system is experimentally characterized, and a total net line rate of 8.4 Gb/s (after FEC overhead subtraction) is demonstrated at a transmission distance of 3200 km.

The Chapter is organized as follows. Firstly, the discussion about TX and RX digital signal processing (DSP) algorithms and the achievable digital B2B performance are reported in Section 6.1. The experimental setup is described in Section 6.2, together with the characterization of the optical B2B performance. The transmission results are reported and discussed in Section 6.3. Finally, the conclusions are drawn in Section 6.4.

This Chapter reports some of the results presented in [26].

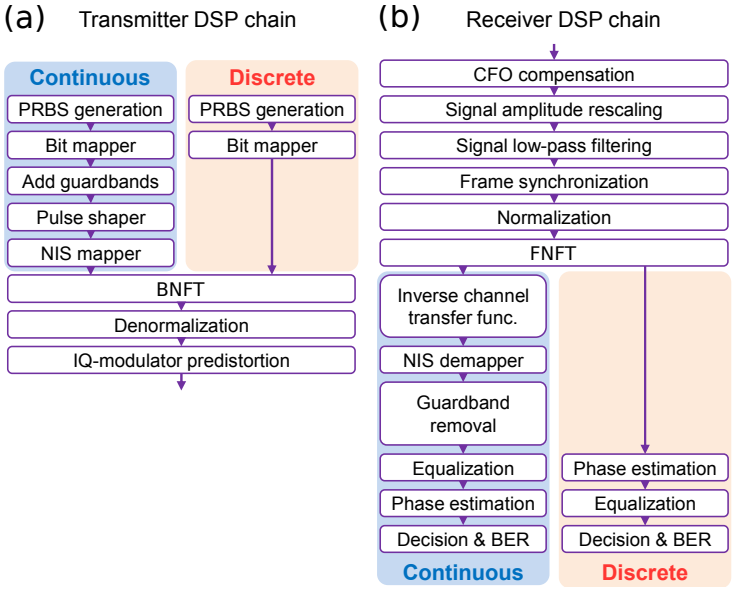


Figure 6.1: DSP chain, highlighting the key operations performed on the digital waveforms at the (a) TX and (b) RX.

6.1 Digital signal processing

The DSP chain implemented at the TX and RX to properly encode the data into a digital waveform and extract it back are shown in Figs. 6.1(a)-(b), respectively.

As far as it concerns the TX, a pseudo-random bit sequence (PRBS) is generated at the TX side to be encoded in both the continuous and discrete spectra. On the one hand for the continuous spectrum, after mapping the bits into 16 10 GBd QPSK symbols (with 16 samples/symbol), guardbands of 64 symbols are added for each 16-symbol burst, leading to an overall burst duration of length of 8 ns. Such guard intervals ensure that no inter-burst interference takes place after the dispersion-induced pulse broadening at the maximum transmission distance considered in this work, i.e., 3200 km. Such long guard intervals could be decreased by adding dispersion precompensation at the transmitter side. This was avoided for the experimental investigation since dispersion precompensation would yield digital waveforms with different peak-to-average power ratio (PAPR) for each transmission distance, impacting the comparison. The symbols are pulse-shaped with a raised cosine (RC) filter (with roll-off equal to 1) and the NIS mapper is used to obtain the continuous spectrum as described in Section 4.2 and detailed in Fig. 6.2. On the other hand, the data bits to be encoded onto the discrete spectrum are mapped to QPSK symbols and then associated to two discrete components. The discrete eigenvalues—two purely-imaginary eigenvalues $\lambda_1 = 0.3j$ and $\lambda_2 = 0.6j$, with time normalization parameter $T_0 = 244$ ps—are not modulated, and both are transmitted in each burst.

The corresponding discrete components are modulated using **b**-modulation[1] for its superior performance [72], i.e., the QPSK symbols are mapped directly on the $\mathbf{b}(\lambda_i)$, rather than on the norming constant $\mathbf{C}_i$, for $i = 1, 2$. The radii of the two QPSK constellations have been set to $5\sqrt{2}$ and $0.05\sqrt{2}$, for λ_1 and λ_2, respectively. This choice leads to a temporal separation of the components of the time-domain signal associated to the NFT coefficients $\mathbf{b}(\lambda_i)$ corresponding to different eigenvalues. The separation is such that the discrete spectral components (at the transmitter output) are placed in time within the guardbands of the continuous-spectrum burst. The one corresponding to λ_1 can be seen in Fig. 6.3(a) after the burst encoded in the continuous spectrum, whereas the waveform corresponding to λ_2 is located on the opposite end of the symbol slot. This choice was made to avoid problems due to high PAPR at the transmitter side as well as to limit the time-frequency product of the multisoliton signal [100]. Nevertheless, continuous and discrete time-components do interact during fiber propagation. A BNFT operation is then performed as described in Section 3.4.3 to generate a time-domain waveform with the desired continuous and discrete spectra. After proper denormalization, the waveform shown in Fig. 6.3(a) is obtained (signal power of -9.2dBm with the fiber parameters of Section 6.2). The figure shows the two discrete (solitonic) components with the continuous (dispersive) components in between. Finally, the waveforms are pre-distorted to account for the nonlinear transfer function of the in-phase and quadrature (IQ) modulator by applying an $\arcsin(\cdot)$ function. Such a digital waveform can then be encoded onto an optical carrier using a standard IQ modulator, after digital-to-analog conversion, as will be described in Section 6.2. The net line rate of the generated signal is 8.4Gb/s, taking into account the 80%-guardbands applied and the 7% hard decision forward error correction (HD-FEC) overhead [101,102].

At the RX side, the digital waveforms are processed by the DSP highlighted in Fig. 6.1(b). First, carrier offset compensation is performed to remove any frequency shift due to frequency mismatch between signal and local oscillator (LO), followed by signal amplitude rescaling, low-pass filtering at twice the 20dB signal bandwidth, and frame synchronization. The FNFT is then applied to recover the continuous and discrete spectra from the time-domain waveform—the trapezoidal discretization method was used [68], using the forward-backward procedure on the discrete spectrum [67, 68]. Concerning the detection of the continuous spectrum, the inverse transfer function of the channel is applied, NIS mapping is inverted, and the guardbands are removed. Next, blind radius-directed equalization is performed followed by phase estimation using digital phase-lock loop. Finally, decisions on the symbols are taken and the BER is counted. On the other hand, the demodulation of the discrete spectrum consists of phase recovery using blind phase search (BPS) independently on each constellation, followed by NFT-domain equalization [72, 88]. As the chosen eigenvalues are purely imaginary, BPS inherently applies the ideal inverse channel transfer function, which consists of a constant phase rotation. After BPS,

[1] Despite the same name, here **b**-modulation is used for discrete components, rather than for the continuous spectrum as in Section 4.6.

NFT-domain equalization reduces the noise on the $\mathbf{b}(\lambda_i)$ by exploiting the correlation between the received eigenvalues and the spectral amplitudes [72, 88]. This equalizer enables to partially compensate for the rotation and re-scaling experienced by $b_1(\lambda_i)$ and $b_2(\lambda_i)$ due to the displacement of the eigenvalues. After equalization, decisions are taken and BER counting is performed.

The DSP chains and numerical algorithms have been first benchmarked in a digital B2B scenario where the digital waveforms before the IQ-modulator predistortion are fed directly into the receiver DSP chain. This analysis allows to focus on the numerical algorithms, while ignoring the impact of practical equipment limitations, such as ADC and DAC resolution, and electrical/optical noise sources. The resulting performance is shown in Fig. 6.3(b), as a function of the energy in the continuous spectrum, comparing joint and continuous-only modulation. For the joint modulation (top figure), the energy in the discrete spectrum is kept constant to fulfill the duration-amplitude relation [11]. The signal quality is evaluated by calculating the EVM, since the BER values are too low for reliable error counting. In the case of joint modulation, as the energy in the continuous spectrum increases, the performance of the continuous spectrum improves (EVM decreases) with an optimum at approximately 0.18pJ (-9.2dBm of launch power). When the energy in the continuous spectrum approaches zero, the discrete spectrum is dominant and worsens the accuracy of the numerical algorithms for the continuous part. In the case of continuous-only modulation (bottom figure), the performance does not degrade as the energy decreases as for the joint-modulation, thus ruling out numerical errors of the NCG alone. Beyond the optimum energy for the joint-modulation, the performance worsens rather rapidly above -9.0dBm. A similar worsening of the performance is indeed reflected when no discrete spectrum is present. Note that, considering **b**-modulation also for the continuous spectrum (see Section 4.6) may provide further improvement even though using it in the context of joint spectral modulation may present some challenges [73]. Figure 6.3(b) shows the impact of the energy in the continuous spectrum on the quality of the discrete spectrum: as the energy in the continuous spectrum is increased, the limited precision of the numerical algorithms yields a loss of orthogonality between continuous and discrete spectrum, thus decreasing the performance of the latter when the continuous components at high energy overlap with the solitons. The impact of the time-overlap is expected to worsen the performance loss during transmission over a non-ideal (lossy and noisy) channel, as discussed in [103]. Furthermore, the numerical precision of the BNFT is also expected to contribute to the overall worsening of the performance of the discrete spectrum. Regardless of these limitations and the consequent energy balance between continuous and discrete spectrum, the BER that can be estimated from the EVM is well below 1×10^{-4}, even for the highest power values considered.

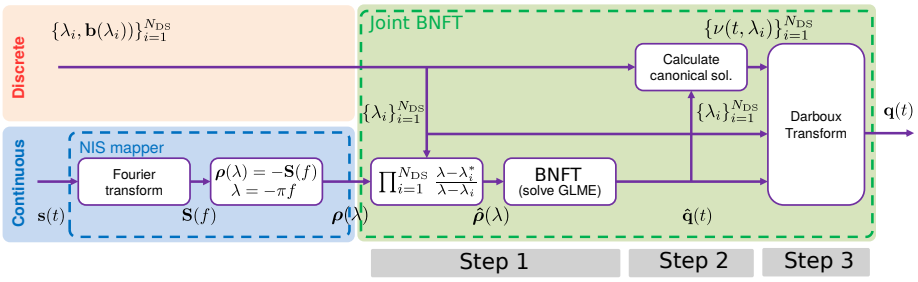


Figure 6.2: Dual-polarization join NFDM signal generation at the TX using a joint NFT operation.

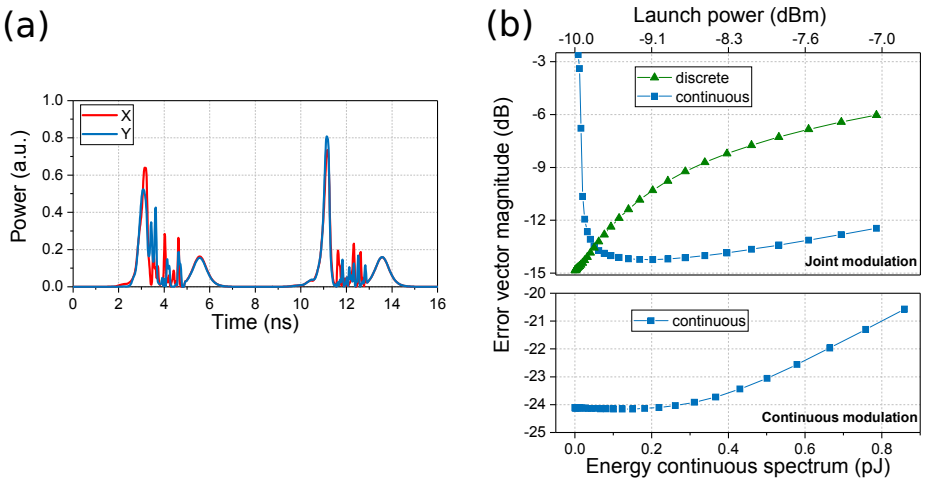


Figure 6.3: (a) Time domain waveform (−9.2dBm launch power, 2 bursts) showing the discrete (solitonic) components with the continuous components in between and (b) digital B2B performance: EVM vs energy in the continuous spectrum for joint modulation (top) and continuous-only modulation (bottom).

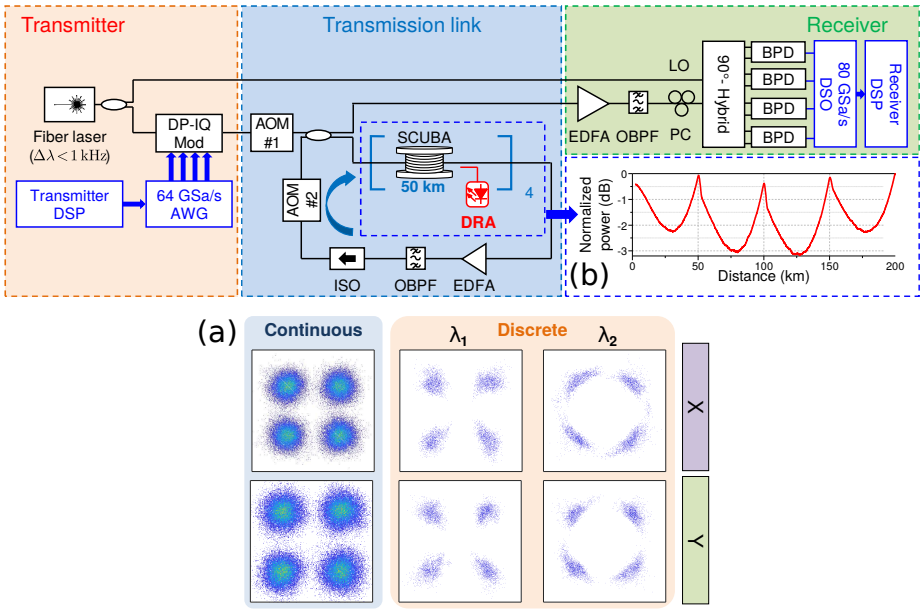


Figure 6.4: Experimental setup for the transmission of the jointly modulated signal in a recirculating transmission loop. Insets: (a) power profile measured over a loop recirculation and (b) constellation diagrams after 2800-km transmission.

6.2 Experimental transmission setup

The experimental setup is shown in Fig. 6.4. The predistorted digital waveforms generated as in Fig. 6.1(a) are loaded into a 4-channel 64-GSamples/s arbitrary waveform generator (AWG) driving the IQ modulator, which encodes the dual-polarization NFDM signal into an optical carrier generated by a low-linewidth ($\leq 1\,\mathrm{kHz}$) fiber laser. The same laser is used as LO at the receiver side.

The transmission link consists of a recirculating transmission loop based on four 50 km transmission spans with distributed Raman amplification applied to each span as in [88]. Backward pumping combined with low-loss large effective area fiber (SCUBA fiber) enables to achieve maximum power variations of approximately 3 dB across the full 200 km loop length. The power profile measured by optical time domain reflectometry is shown in inset (a) of Fig. 6.4. Loss, dispersion, and nonlinear coefficient of the transmission fiber are $0.155\,\mathrm{dB/km}$, $22\,\mathrm{ps/nm/km}$, and $\gamma = 0.6\,\mathrm{W^{-1}km^{-1}}$, respectively. These values have been used for the NFT normalization as discussed in Section 3.2.1. In addiction to the transmission fiber, the loop consists of acusto-optic modulator (AOM) used as optical switches, an optical band pass filter (OBPF) (0.5 nm bandwidth) which suppresses out-of-band ASE noise, an isolator (ISO), and an erbium-doped fiber amplifier (EDFA) which compensates for the power loss of all these components.

After the chosen number of recirculation turns, the signal is received with a pre-amplified coherent receiver using four balanced photodetectors (BPDs) and a 80-GSamples/s digital storage oscilloscope (DSO) acting as ADC. For simplicity, the signal polarization is manually aligned at the receiver input with a polarization controller (PC). However, demultiplexing schemes based on training sequences have already been reported [87]. After analog-to-digital conversion, the waveforms are processed offline by the DSP discussed in Section 6.1 and the performance are evaluated by bit error counting performed on more than 10^6 bits, ensuring a reliable BER above 10^{-5}. In the following, the transmission reach is evaluated considering the HD-FEC threshold (BER of 3.8×10^{-3}) [101,102]. We remark that the frequency-offset estimation discussed in Section 6.1 is necessary due to the frequency shift introduced by the AOM which results in self-heterodyne detection rather than homodyne. An example of constellation diagrams for continuous and discrete spectrum is shown in inset (b) of Fig. 6.4 after 2800 km transmission, illustrating the high quality of the received signals.

Before discussing the transmission results in Section 6.3, the signal performance are evaluated in B2B configuration, i.e., connecting the receiver directly at the output of the IQ modulator. The results are shown in Fig. 6.5, distinguishing between the performance of continuous and discrete spectral components as well as showing the total BER. The optical signal-to-noise ratio (OSNR) at the output of the transmitter was approximately 33.8 dB for all the launch powers considered.

The total BER is dominated by the BER of the continuous spectrum as more bits are encoded in the continuous spectrum compared to the discrete spectrum

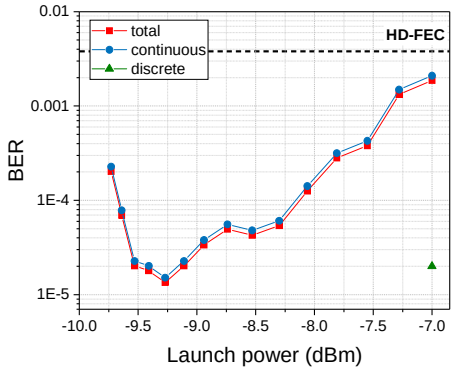


Figure 6.5: Optical B2B BER performance as a function of the launch signal power.

for each NFDM symbol. As discussed for the numerical results of Fig. 6.3, the BER of the continuous spectrum improves with its increased energy, reaches an optimum and worsens due to numerical instabilities. Note that the optimum power is the same for the digital B2B (see Fig. 6.3(b)). The lack of variations in the optimum power hints that the dominant limitation is currently related to the numerical algorithms, whereas the impact of electrical/optical noise at the TX and the RX, as well as the AWG resolution are rather negligible. The BER on the discrete spectrum is also consistent with Fig. 6.3(b), as errors are only detected at the highest launch power considered, -7 dBm. At such a power level, the estimated BER in digital B2B was estimated to almost the same value, showing that a negligible penalty is introduced by the optical-frontends (at both transmitter and receiver) also for the discrete spectrum.

6.3 Transmission performance

This Section reports the transmission performance of the dual-polarization NFDM system described in the previous Sections; the digital and optical B2B performance have also been discussed. Figure 6.6 shows the evolution along the fiber of one burst at the optimum launch power of -9.2 dBm. Firstly, the waveforms show that the two spectral components—the continuous and the discrete part—interact in time; secondly, the figure shows that the guardband size is more than sufficient to guarantee the vanishing boundary conditions required by the NFT also at the longest transmission distances. The guardband size could actually be reduced by pre-dispersing the waveforms at the transmitter side by half of the transmission length, as in Section 4.4 [75,85]. Additionally, by tailoring the guardbands to the desired transmission distance, the transmission rate can be maximized.

Figure 6.7(a) shows the total BER as a function of the launch power for different transmission distances. The curves show an optimum launch power

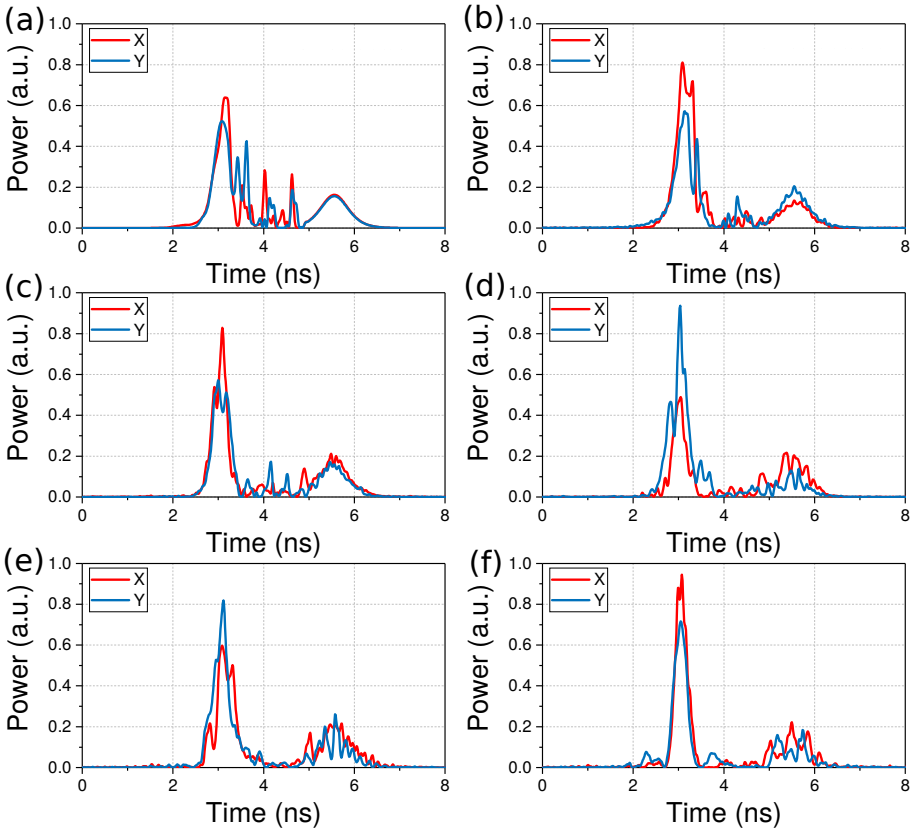


Figure 6.6: Examples of time-domain waveforms showing one 8-ns burst (including guardbands) at a fixed launch power of −9.2 dBm: (a) digital B2B, (b) optical B2B, (c) 400 km, (d) 1200 km, (e) 2000 km, and (f) 2800 km transmission. The bit patter is not the same for the different waveforms.

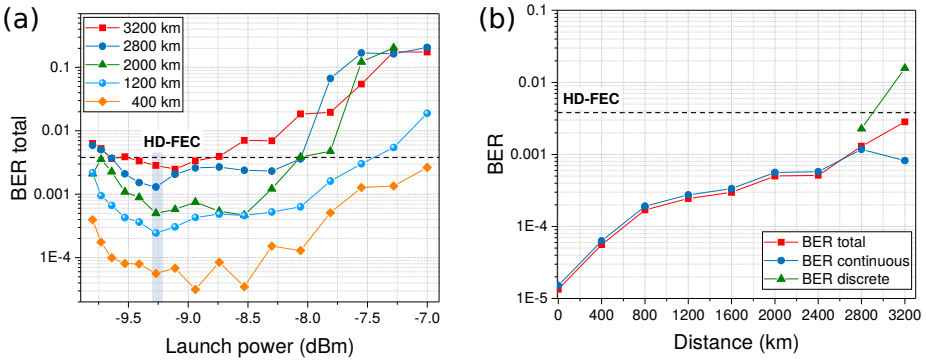


Figure 6.7: (a) Total BER performance as a function of the launch power for different transmission distances, and (b) total BER and contributions from continuous and discrete spectrum as a function of the transmission distance for the optimum launch power of -9.2 dBm (shaded area in (a)).}

(minimum BER) consistent with the digital and optical B2B performance, i.e., -9.2 dBm. The BER values after a 400 km transmission are actually in close agreement with the optical B2B ones in Fig. 6.5. These results further confirm that the dominant performance limitation is currently linked to the numerical precision in performing the NFT operations at high power values. By increasing the numerical precision of the different NFT operations and/or using improved numerical algorithms, we believe the performance could be significantly improved, potentially shifting the optimum launch power to higher values.

The BER as a function of the transmission distance is shown in Fig. 6.7(b), at the fixed launch power of -9.2 dBm. The figure highlights that BER below the HD-FEC threshold can be achieved for up to 3200 km transmission. Beyond 2800 km, the dominant contribution to the total BER comes from the discrete spectral components and in particular from the largest eigenvalue $\lambda_2 = 0.6j$.

6.4 Conclusion and outlook

This Chapter introduced a framework for dual-polarization NFDM systems which allows to encode data on both continuous and discrete spectral components. The steps to perform the joint BNFT, described in Section 3.4.3, are verified numerically, in a transmission-free scenario without (digital B2B) and with the optical front-ends (optical B2B). The dual-polarization joint NFDM system has then been experimentally demonstrated in a transmission scenario using distributed Raman amplification. A transmission reach of 3200 km is achieved for a 8.4 Gb/s net rate NFDM signal, mainly limited by the numerical implementation of the NFT, which will need to be further improved. This work demonstrates for the first time the use of all the degrees of freedom available

for NFDM-based transmission over SMFs.

As a future work, we expect to improve the performance of the dual-polarization NFDM with joint spectrum in different ways. Firstly, considering different modulation and demodulation schemes and improved detection strategies for the continuous spectrum [54, 86] (see Section 4.5), and/or differential modulation [104] or soliton detection based on matched filter [14] for the discrete spectrum. Secondly, the optimization of the discrete eigenvalues parameters (energy, velocity, and position) with respect to the continuous spectrum and the overall transmission rate deserves investigation and, we believed, can provide some performance improvement. Thirdly, improvements in the numerical algorithms are essential to improve the reach of the scheme, as well as reduce the computational cost. Finally, the net transmission rate may be increased by reducing the guard bands, as mentioned above, and predispersing the signal by half of the transmission length at the transmitter side, as in Section 4.4.

Chapter 7

Conclusion and outlook

This thesis considered NFDM transmission systems—i.e., communication systems in which information is encoded on the nonlinear spectrum using the NFT—for optical fiber communications, aiming at improving the performance of the current optical network. Such systems have been investigated with a critical approach and driven by a medium-long term, rather than immediate, perspective. The work has been carried out through different steps, starting from the mathematical and theoretical framework, going through the development of the numerical algorithms, the investigation of system performance with simulations, and, finally, some experimental demonstrations.

The main contributions and outcomes of the work are discussed in the following.

- The mathematical framework underlying the NFDM transmission technique has been reported. In particular, the theoretical aspects and properties relevant for optical communication purposes of the NFT regarding the VNLSE with vanishing boundary conditions have been discussed. Some properties have been proved.

- An algorithmic framework for the NFT operations regarding the VNLSE with vanishing boundary conditions—essential for the realization of the corresponding NFDM system—has been developed. In particular, the numerical algorithms for the multidimensional case were derived by extending those conceived for the NLSE case. Concerning the FNFT, the Boffetta–Osborne method was extended to the VNLSE with two versions: the exact method, and an approximated, but less complex, one. Regarding the BNFT, the NCG method—a method to solve the BNFT via the GLME—for the VNLSE was developed and described in details; to the best of our knowledge, it is the only method for the solution of the GLME for $M \geq 2$, and can, in principle, include both spectra. Moreover, a methodology based on the DT for the modulation of both spectra for the ME case has been derived, following that for the scalar NLSE. All the algorithms have been tested to understand—and, possibly, exclude—their

impact on the system performance. Note, however, that the optimization of the algorithms was out of the scope of the work. A reduction of their complexity is, therefore, a possible subject of future investigation.

- Single-polarization NFDM systems have been studied. In particular, the NIS framework with only continuous spectrum modulation was considered. The performance of such system was investigated through simulations and compared with theoretical estimations and conventional systems. Some critical aspects of NFDM were highlighted and discussed, in particular concerning the detrimental noise dependence (in the nonlinear spectrum) on the spectrum itself. This fact cause the worsening of the system performance when increasing the rate efficiency, which makes conventional systems superior. We show that this behavior is due to the sub-optimal detection strategy considered at the RX, indeed tailored for linear systems. In the light of the above, and aiming at developing an optimal detection strategy for NFDM (over the AWGN channel), we proposed some improved detection strategies, and we showed that significant performance improvements can be obtained (up to 6.2 dB in Q^2) with DF-BNFT detection. Furthermore, we discussed other strategies aimed at mitigate NFDM issues, such as precompensation, and some recently proposed modulation schemes such as **b**-modulation. In conclusion, we showed that (i) commonly considered NFDM schemes can not compete with conventional systems, because a non-tailored detection strategy inherited from linear system is used, and (ii) tailored detection strategies can reveal the full potential of NFDM and, hopefully, outperform conventional systems.

- Dual-polarization systems using the NIS framework have been discussed. We showed that using this scheme, we can double the information carried with a slight performance worsening with respect to a single-polarization scheme. We expect that the concepts discussed for single-polarization (noise impact on the nonlinear spectrum) still hold and that similar modulation and demodulation techniques can be used. Moreover, we show that a reduced complexity scheme, tough introducing a deterministic mismatch with the channel, may provide some performance improvement.

- The first experimental demonstration a of dual-polarization NFDM system encoding information on both the discrete and the continuous spectrum has been described and discussed. A net data rate of 8.4 Gb/s at a transmission distance of 3200 km on a SMF using Raman amplification was obtained encoding 4 QPSK symbols on two discrete eigenvalues (**b**-modulation, 2 symbols per polarization), and 32 QPSK symbols on the continuous spectrum (NIS mapping, 16 symbols per polarization).

The results obtained in this thesis, together with several excellent works carried out by other research groups in the last five years, are a first important step towards understanding NFDM systems. Some fundamental issues have been investigated. However, the recent results showed that the topic is not straightforward, and a deeper study and tailored design are necessary, which may not

be obtained in a few years. Nevertheless, also the communication systems used nowadays have not been developed in a few years.

There are still many challenges and interesting subjects to be investigated about NFDM, some already mentioned throughout the work. Some of these are reported in the following.

- The large guard times—carrying no information—required for the NFT boundary conditions and to avoid burst-burst overlapping due to dispersion may be reduced considering a system based on the NFT with periodic boundary conditions [17,52]. In this way, considering a cyclic prefix rather than long guard times, the spectral efficiency (SE) may significantly increase.

- A further optimization of the modulation and demodulation processing in the simplest scenario—single-user, single-polarization, AWGN lossless channel (or even B2B)—is essential to enable the success of NFDM systems. This include the understanding of the statistics of noise in the nonlinear spectrum, as well as the use of modulation and demodulation processing tailored for NFDM.

- The development of fast and accurate numerical algorithms for the NFT (also for the $M \geq 2$ case) is essential for the success of NFDM. This would also allow to use the NFT for DBP.

- The development of optical FNFT and BNFT [81] is crucial to effectively assign different nonlinear frequencies to different users. In this way one could replace traditional ROADM with NADM—the name was introduced in [11]—to multiplex different users.

- It is worth investigating about the possibility of SDM-NFDM schemes; such schemes could take into account deterministic mode/core interactions.

Appendix A

Additional mathematical definitions, proofs, and properties

A.1 Complex envelop of real signals

A real bandpass signal $x(t)$ with bandwidth B whose frequencies are centered around $\pm f_0$, is a signal whose FT $\mathcal{F}\{x(t)\}(f)$ is null outside $[\pm f_0 - B/2, \pm f_0 + B/2]$ and symmetric with respect to the zero frequency. When $B \ll 2f_0$, $s(t)$ is said to be a narrow bandpass signal. The bandpass signal $x(t)$ can be represented as

$$x(t) = A(t)\cos(2\pi f_0 + \phi(t)), \tag{A.1}$$

with $A(t)$ and $\phi(t)$ being the envelope and phase, respectively. The complex envelop, or lowpass equivalent, of $x(t)$ is the complex signal [39]

$$w(t) = A(t)e^{j\phi(t)} = z(t)e^{-j2\pi f_0 t} = i(t) + jq(t) \tag{A.2}$$

where the analytic signal $z(t)$ is

$$z(t) = x(t) + j\check{x}(t), \tag{A.3}$$

$\check{x}(t)$ being the Hilbert transform of $x(t)$. The analytic signal has only positive frequency components. Importantly, the complex envelop $w(t)$ is a complex lowpass signal with bandwidth B, i.e., its FT is null outside $[-B/2, B/2]$. The in-phase and quadrature components—$i(t)$ and $q(t)$, respectively—are the real and imaginary part of $w(t)$ [39]. The bandpass signal can be obtained from its complex envelop as

$$x(t) = \Re\left\{w(t)e^{2\pi j f_0 t}\right\} = \left[w(t)e^{2\pi j f_0 t} + w(t)^* e^{-2\pi j f_0 t}\right]/2. \tag{A.4}$$

It is useful to work with the complex envelop $w(t)$, rather than the real signal $x(t)$, to account for the slowly varying (with respect to the carrier frequency f_0) envelope and phase.

A.2 Transfer matrix for the numerical forward NFT in Section 3.4.1

In this Section, we derive an exact form for the transfer matrix $U^{(n)} = \exp\left(P^{(n)}\delta\right)$ appearing in Eq. (3.78), with $P^{(n)}$ defined in Eq. (3.73). To simplify the notation, let us drop the superscript $^{(n)}$. The matrix $P^{(n)}$ (of dimension $(M+1)\times(M+1)$) can be diagonalized as $P^{(n)} = O\Lambda O^{-1}$ where

$$\Lambda = \begin{pmatrix} d & & & & & \\ & -d & & & & \\ & & j\lambda & & & \\ & & & \ddots & & \\ & & & & j\lambda & \end{pmatrix}, \tag{A.5}$$

$d = \sqrt{-\lambda^2 - \sigma\sum_{k=1}^{M}|q_k|^2}$, and

$$O = \begin{pmatrix} d-j\lambda & -(d+j\lambda) & 0 & \dots & \dots & 0 \\ -\sigma q_1^* & -\sigma q_1^* & -q_2 & 0 & \dots & 0 \\ \vdots & \vdots & q_1 & -q_3 & \dots & \vdots \\ \vdots & \vdots & 0 & q_2 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & -q_M \\ -\sigma q_M^* & -\sigma q_M^* & 0 & 0 & 0 & q_{M-1} \end{pmatrix}. \tag{A.6}$$

Equivalently, the (m,ℓ)-entry of O is

$$O_{m,\ell} = \begin{cases} (d-j\lambda) & m=\ell=1 \\ -(d+j\lambda) & m=1,\ \ell=2 \\ 0 & m=1,2,\ \ell\geq m+2 \\ -\sigma q_{m-1}^* & \ell=1,2,\ m\geq 2 \\ q_{m-2} & m=\ell,\ \ell\geq 3 \\ -q_m & m=\ell+1,\ \ell\geq 3 \\ 0 & m,l\geq 3,\ m\neq \ell,\ell-1 \end{cases}. \tag{A.7}$$

Moreover, the (m,ℓ) component of O^{-1}, the inverse of O, is

$$O^{-1}_{m,\ell} = \frac{1}{2d\sum_{k=1}^{M}|q_k|^2}\begin{cases}\sum_{k=1}^{M}|q_k|^2 & m=\ell=1\\ -\sum_{k=1}^{M}|q_k|^2 & m=2,\ \ell=1\\ 0 & \ell=1,\ m\geq 3\\ -\sigma(d+j\lambda)q_{\ell-1} & m=1,\ \ell\geq 2\\ -\sigma(d-j\lambda)q_{\ell-1} & m=2,\ \ell\geq 2\\ \frac{2dq_{\ell-1}}{q_{m-1}q_{m-2}}\sum_{k\leq m-2}|q_k|^2 & 3\leq m\leq l,\ \ell\geq 3\\ -\frac{2dq_{\ell-1}}{q_{m-1}q_{m-2}}\sum_{k\geq m-1}|q_k|^2 & m\geq \ell+1,\ \ell\geq 2\end{cases}. \quad \text{(A.8)}$$

Indeed, carrying out the calculations, one can prove that $P^{(n)} = O\Lambda O^{-1}$. The calculations for $M=2$ are reported in the following:

$$\begin{aligned} O\Lambda O^{-1} &= \left[2d(|q_1|^2+|q_2|^2)\right]^{-1}\begin{pmatrix} d-j\lambda & -d-j\lambda & 0\\ -\sigma q_1^* & -\sigma q_1^* & -q_2\\ -\sigma q_2^* & -\sigma q_2^* & q_1\end{pmatrix}\times\\ &\times\begin{pmatrix} d(|q_1|^2+|q_2|^2) & -\sigma d(d+j\lambda)q_1 & -\sigma d(d+j\lambda)q_2\\ d(|q_1|^2+|q_2|^2) & \sigma d(d-j\lambda)q_1 & \sigma d(d-j\lambda)q_2\\ 0 & -2j\lambda dq_2^* & 2j\lambda dq_1^*\end{pmatrix}\end{aligned} \quad \text{(A.9)}$$

and

$$\begin{aligned}
\left[O\Lambda O^{-1}\right]_{11} &= \left[2d(|q_1|^2+|q_2|^2)\right]^{-1}\left[d(|q_1|^2+|q_2|^2)(-2j\lambda)\right] = -j\lambda = P_{11}^{(n)}\\
\left[O\Lambda O^{-1}\right]_{12} &= \left[2d(|q_1|^2+|q_2|^2)\right]^{-1}\sigma dq_1\left[-2(d^2+\lambda^2)\right] = q_1 = P_{12}^{(n)}\\
\left[O\Lambda O^{-1}\right]_{13} &= \left[2d(|q_1|^2+|q_2|^2)\right]^{-1}\sigma dq_2\left[-2(d^2+\lambda^2)\right] = q_2 = P_{13}^{(n)}\\
\left[O\Lambda O^{-1}\right]_{21} &= \left[2d(|q_1|^2+|q_2|^2)\right]^{-1}\left[-2\sigma q_1^* d(|q_1|^2+|q_2|^2)\right] = -\sigma q_1^* = P_{21}^{(n)}\\
\left[O\Lambda O^{-1}\right]_{22} &= \left[2d(|q_1|^2+|q_2|^2)\right]^{-1}\left[2j\lambda d(|q_1|^2+|q_2|^2)\right] = j\lambda = P_{22}^{(n)}\\
\left[O\Lambda O^{-1}\right]_{23} &= \left[2d(|q_1|^2+|q_2|^2)\right]^{-1}\left[2j\lambda dq_2q_1^* - 2j\lambda dq_1^*q_2\right] = 0 = P_{23}^{(n)}\\
\left[O\Lambda O^{-1}\right]_{31} &= \left[2d(|q_1|^2+|q_2|^2)\right]^{-1}\left[-2\sigma q_2^* d(|q_1|^2+|q_2|^2)\right] = -\sigma q_2^* = P_{31}^{(n)}\\
\left[O\Lambda O^{-1}\right]_{32} &= \left[2d(|q_1|^2+|q_2|^2)\right]^{-1}\left[dq_1q_2^*2j\lambda - 2j\lambda dq_2^*q_1\right] = 0 = P_{32}^{(n)}\\
\left[O\Lambda O^{-1}\right]_{33} &= \left[2d(|q_1|^2+|q_2|^2)\right]^{-1}\left[2j\lambda d(|q_1|^2+|q_2|^2)\right] = j\lambda = P_{33}^{(n)}
\end{aligned}. \quad \text{(A.10)}$$

Consequently,

$$U^{(n)} = \exp\left(P^{(n)}\delta\right) = \exp\left(O\Lambda O^{-1}\delta\right) = O\exp\left(\Lambda\delta\right)O^{-1} = \quad \text{(A.11)}$$

$$= O\begin{pmatrix} e^{d\delta} & & & & \\ & e^{-d\delta} & & & \\ & & e^{j\lambda\delta} & & \\ & & & \ddots & \\ & & & & e^{j\lambda\delta}\end{pmatrix}O^{-1} \quad \text{(A.12)}$$

and one obtains

$$(U^{(n)})_{m,\ell} = \begin{cases} c - j\lambda s & m = \ell = 1 \\ q_{\ell-1} s & m = 1,\ \ell \geq 2 \\ -\sigma q^*_{m-1} s & m \geq 2,\ \ell = 1 \\ r_{m-1,\ell-1}\left[c + j\lambda s - e^{j\lambda\delta}\right] & m = 2,\ \ell \geq 3 \text{ or } \ell = 2,\ m \geq 3 \\ r_{m-1,\ell-1}\left[c + j\lambda s\right] + e^{j\lambda\delta}(1 - r_{m-1,m-1}) & m = \ell = 2 \text{ or } m, \ell \geq 3 \end{cases} \tag{A.13}$$

where $r_{m\ell} = q^*_m q_\ell / \sum_{k=1}^{M} |q_k|^2$, $c = \cosh(d\delta)$, and $s = \sinh(d\delta)/d$. Again, this equation can be verified carrying out the calculations, which are reported in the following for $M = 2$:

$$\begin{aligned} U^{(n)} &= O \exp(\Lambda\delta) O^{-1} = \\ &= \left[2d(|q_1|^2 + |q_2|^2)\right]^{-1} \begin{pmatrix} d - j\lambda & -d - j\lambda & 0 \\ -\sigma q_1^* & -\sigma q_1^* & -q_2 \\ -\sigma q_2^* & -\sigma q_2^* & q_1 \end{pmatrix} \times \\ &\times \begin{pmatrix} (|q_1|^2 + |q_2|^2) e^{d\delta} & -\sigma(d + j\lambda) q_1 e^{d\delta} & -\sigma(d + j\lambda) q_2 e^{d\delta} \\ -(|q_1|^2 + |q_2|^2) e^{-d\delta} & -\sigma(d - j\lambda) q_1 e^{-d\delta} & -\sigma(d - j\lambda) q_2 e^{-d\delta} \\ 0 & -2d q_2^* e^{j\lambda\delta} & 2d q_1^* e^{j\lambda\delta} \end{pmatrix} \end{aligned} \tag{A.14}$$

and

$$\begin{aligned} U_{11}^{(n)} &= [2d]^{-1}\left[(d - j\lambda)e^{d\delta} + (d + j\lambda)e^{-d\delta}\right] = [(e^{d\delta} + e^{-d\delta})/2 + \\ &-j\lambda(e^{d\delta} - e^{-d\delta})/2d] = [\cosh(d\delta) - j\lambda \sinh(d\delta)/d]\,; \end{aligned} \tag{A.15}$$

$$\begin{aligned} U_{12}^{(n)} &= \left[2d(|q_1|^2 + |q_2|^2)\right]^{-1} (-\sigma q_1)(d^2 + \lambda^2)\left[e^{d\delta} - e^{-d\delta}\right] = \\ &= q_1\left(e^{d\delta} - e^{-d\delta}\right)/2d = q_1 \sinh(d\delta)/d; \end{aligned} \tag{A.16}$$

$$\begin{aligned} U_{13}^{(n)} &= \left[2d(|q_1|^2 + |q_2|^2)\right]^{-1} (-\sigma q_2)(d^2 + \lambda^2)\left[e^{d\delta} - e^{-d\delta}\right] = \\ &= q_2\left(e^{d\delta} - e^{-d\delta}\right)/2d = q_2 \sinh(d\delta)/d; \end{aligned} \tag{A.17}$$

$$\begin{aligned} U_{21}^{(n)} &= \left[2d(|q_1|^2 + |q_2|^2)\right]^{-1} (-\sigma q_1^*)(|q_1|^2 + |q_2|^2)\left[e^{d\delta} - e^{-d\delta}\right] = \\ &= -\sigma q_1^*\left(e^{d\delta} - e^{-d\delta}\right)/2d = -\sigma q_1^* \sinh(d\delta)/d; \end{aligned} \tag{A.18}$$

$$\begin{aligned} U_{22}^{(n)} &= \left[2d(|q_1|^2 + |q_2|^2)\right]^{-1} \{|q_1|^2[d(e^{d\delta} + e^{-d\delta}) + j\lambda(e^{d\delta} - e^{-d\delta})] + \\ &+2d|q_2|^2 e^{j\lambda\delta}\} = r_{1,1}\left[\cosh(d\delta) + j\lambda \sinh(d\delta)/d\right] + r_{2,2} e^{j\lambda\delta}; \end{aligned} \tag{A.19}$$

$$\begin{aligned} U_{23}^{(n)} &= \left[2d(|q_1|^2 + |q_2|^2)\right]^{-1} \{q_1^* q_2[d(e^{d\delta} + e^{-d\delta}) + j\lambda(e^{d\delta} - e^{-d\delta})] + \\ &-2d q_2 q_1^* e^{j\lambda\delta}\} = r_{1,2}\left[\cosh(d\delta) + j\lambda \sinh(d\delta)/d - e^{j\lambda\delta}\right]; \end{aligned} \tag{A.20}$$

$$\begin{aligned} U_{31}^{(n)} &= \left[2d(|q_1|^2 + |q_2|^2)\right]^{-1} (-\sigma q_2^*)(|q_1|^2 + |q_2|^2)\left[e^{d\delta} - e^{-d\delta}\right] = \\ &= -\sigma q_2^*\left(e^{d\delta} - e^{-d\delta}\right)/2d = -\sigma q_2^* \sinh(d\delta)/d; \end{aligned} \tag{A.21}$$

$$\begin{aligned} U_{32}^{(n)} &= \left[2d(|q_1|^2 + |q_2|^2)\right]^{-1} \{q_1 q_2^*[d(e^{d\delta} + e^{-d\delta}) + j\lambda(e^{d\delta} - e^{-d\delta})] + \\ &-2d q_1 q_2^* e^{j\lambda\delta}\} = r_{2,1}\left[\cosh(d\delta) + j\lambda \sinh(d\delta)/d - e^{j\lambda\delta}\right]; \end{aligned} \tag{A.22}$$

$$U_{33}^{(n)} = \left[2d(|q_1|^2+|q_2|^2)\right]^{-1}\{|q_2|^2[d(e^{d\delta}+e^{-d\delta})+j\lambda(e^{d\delta}-e^{-d\delta})]+ \\ +2d|q_1|^2e^{j\lambda\delta}\} = r_{2,2}\left[\cosh(d\delta)+j\lambda\sinh(d\delta)/d\right]+r_{1,1}e^{j\lambda\delta}. \tag{A.23}$$

The inverse matrix of $U^{(n)}$ can be obtained similarly from $U^{(n)-1} = O\Lambda^{-1}O^{-1}$ and reduces to Eq. (3.90).

A.3 Obtaining the vector Gelfand-Levitan-Marchenko equation

In [9], Chapter 4, pp. 108-109, the authors derive the VGLMEs

$$\bar{\mathbf{K}}(x,y)+\begin{pmatrix}\mathbf{0}_{1\times M}\\ \mathbf{I}_M\end{pmatrix}\mathbf{F}(x+y)+\int_x^\infty \mathbf{K}(x,s)\mathbf{F}(s+y)\,ds=0 \tag{A.24}$$

$$\mathbf{K}(x,y)+\begin{pmatrix}1\\ \mathbf{0}_{M\times 1}\end{pmatrix}\bar{\mathbf{F}}(x+y)+\int_x^\infty \bar{\mathbf{K}}(x,s)\bar{\mathbf{F}}(s+y)\,ds=0 \tag{A.25}$$

for the unknowns $\mathbf{K}(x,y)\in\mathbb{C}^{(1+M)\times M}$ and $\bar{\mathbf{K}}(x,y)\in\mathbb{C}^{(1+M)\times 1}$, given $\mathbf{F}(x)\in\mathbb{C}^{M\times 1}$ and $\bar{\mathbf{F}}(x)\in\mathbb{C}^{1\times M}$ functions of the nonlinear spectrum [9]. Let

$$\mathbf{K}=\begin{pmatrix}\mathbf{K}^{(up)}\\ \mathbf{K}^{(dn)}\end{pmatrix} \text{ and } \bar{\mathbf{K}}=\begin{pmatrix}\bar{\mathbf{K}}^{(up)}\\ \bar{\mathbf{K}}^{(dn)}\end{pmatrix} \tag{A.26}$$

where $\mathbf{K}^{(up)}$ is $1\times M$, $\mathbf{K}^{(dn)}$ is $M\times M$, $\bar{\mathbf{K}}^{(up)}$ is 1×1, and $\bar{\mathbf{K}}^{(dn)}$ is $M\times 1$. The time domain signal is obtained from Eqs. (A.24)-(A.25) as

$$\mathbf{q}(t)=-2\mathbf{K}^{(up)}(t,t)=2\sigma(\bar{\mathbf{K}}^{(dn)}(t,t))^\dagger. \tag{A.27}$$

To obtain Eq. (3.29), rewrite Eqs. (A.24)-(A.25) in terms of $\mathbf{K}^{(up)}$, $\mathbf{K}^{(dn)}$, $\bar{\mathbf{K}}^{(up)}$, and $\bar{\mathbf{K}}^{(dn)}$ obtaining 4 equations

$$\bar{\mathbf{K}}^{(up)}(x,y)+\int_x^\infty \mathbf{K}^{(up)}(x,s)\mathbf{F}(s+y)\,ds=0,\ \text{dimension } 1\times 1 \tag{A.28}$$

$$\bar{\mathbf{K}}^{(dn)}(x,y)+\mathbf{F}(x+y)+\int_x^\infty \mathbf{K}^{(dn)}(x,s)\mathbf{F}(s+y)\,ds=0,\ \text{dimension } M\times 1 \tag{A.29}$$

$$\mathbf{K}^{(up)}(x,y)+\bar{\mathbf{F}}(x+y)+\int_x^\infty \bar{\mathbf{K}}^{(up)}(x,s)\bar{\mathbf{F}}(s+y)\,ds=0,\ \text{dimension } 1\times M \tag{A.30}$$

$$\mathbf{K}^{(dn)}(x,y)+\int_x^\infty \bar{\mathbf{K}}^{(dn)}(x,s)\bar{\mathbf{F}}(s+y)\,ds=0,\ \text{dimension } M\times M. \tag{A.31}$$

Inserting Eq. (A.28) in Eq. (A.30) one obtains

$$\mathbf{K}^{(up)}(x,y)+\bar{\mathbf{F}}(x+y)-\int_x^\infty\int_x^\infty \mathbf{K}^{(up)}(x,r)\mathbf{F}(r+s)\,dr\bar{\mathbf{F}}(s+y)\,ds=0. \tag{A.32}$$

Finally, considering that [9]

$$\bar{\mathbf{F}}(x)=-\sigma\mathbf{F}^\dagger(x) \tag{A.33}$$

and defining $\mathbf{K}=\mathbf{K}^{(up)}$ one obtains the Eq. (3.29).

A.4 Numerical Hankel matrix-vector products

The product of the matrix $\mathcal{H}$ for a column vector $\mathbf{x}$ of length $N_{\mathcal{H}}$ can be performed considering the first $N_{\mathcal{H}}$ components resulting from the product of the circulant matrix of doubled dimensions $\mathcal{C} = \mathcal{C}(\mathbf{h}_0)$ generated by the vector $\mathbf{h}_0 = (\mathbf{h}, \mathbf{0}_{1\times N_{\mathcal{H}}})$ with the vector $\mathbf{x}_0 = (\mathbf{x}^{\mathrm{T}}, \mathbf{0}_{1\times N_{\mathcal{H}}})^{\mathrm{T}}$. Specifically, if $\mathcal{C}(\mathbf{h}_0)\mathbf{x}_0 = \mathbf{y}_0 = (\mathbf{y}^{\mathrm{T}}, \mathbf{w}^{\mathrm{T}})^{\mathrm{T}}$ where both $\mathbf{y}$ and $\mathbf{w}$ are column vector of $N_{\mathcal{H}}$ components, then $\mathbf{y} = \mathcal{H}\mathbf{x}$. The product $\mathbf{y}_0 = \mathcal{C}(\mathbf{h}_0)\mathbf{x}_0$ is the discrete circular convolution between the vectors $\tilde{\mathbf{h}}_0$, where $\tilde{\mathbf{h}}_0^{\mathrm{T}} = (h_1, \mathbf{0}_{1\times N_{\mathcal{H}}}, h_{N_{\mathcal{H}}}, \ldots, h_2)^{\mathrm{T}}$, and $\mathbf{x}_0$, and therefore can be efficiently computed numerically through the FFT operations as IFFT(FFT($\tilde{\mathbf{h}}_0$)·FFT($\mathbf{x}_0$)), where · indicates point to point multiplication.

A.5 Simpson quadrature rule

The composite Simpson's quadrature rule numerically integrates the function $f(x)$ in $[a, b]$ by applying the Simpson's rule on n sub-intervals of width $\Delta = (b-a)/n$. If $x_k = a + (k-1)\Delta$ for $k = 1, \ldots, n+1$, the integral value is approximated as

$$\int_a^b f(x)\,dx \approx \frac{\Delta}{3}\left[f(x_1) + 4\sum_{j=1}^{n/2} f(x_{2j}) + 2\sum_{j=1}^{n/2-1} f(x_{2j+1}) + f(x_{n+1})\right], \quad \text{(A.34)}$$

i.e., the values of $f(x)$ on the grid $x_1, \ldots, x_{n+1}$ of $[a, b]$ are numerically integrated with the weights $\mathbf{d} = \Delta/3(1, 4, 2, 4, \ldots, 4, 2, 4, 1)$.

List of publications

S. Civelli, S. Turitsyn, M. Secondini, and J. Prilepsky, "Polarization-multiplexed nonlinear inverse synthesis with standard and reduced-complexity NFT processing: erratum", Optics Express **27**, 3617–3617 (2019).

F. Da Ros, **S. Civelli**, S. Gaiarin, E. P. da Silva, N. De Renzis, M. Secondini, and D. Zibar, "Dual-polarization NFDM transmission with continuous and discrete spectral modulation," Journal of Lightwave Technology (2019).

S. Civelli, S. K. Turitsyn, M. Secondini, and J. E. Prilepsky, "Reduced complexity nonlinear inverse synthesis transmission method with polarization division multiplexing," in "ECOC 2018; 44nd European Conference on Optical Communication; Proceedings of," We4F.2 (2018).

S. Civelli, E. Forestieri, and M. Secondini, "Improved detection strategies for nonlinear frequency-division multiplexing," arXiv preprint arXiv:1805.02047, in "Progress In Electromagnetics Research Symposium (PIERS) 2018" (2018).

S. Civelli, S. Turitsyn, M. Secondini, and J. Prilepsky, "Polarization-multiplexed nonlinear inverse synthesis with standard and reduced-complexity NFT processing", Optics Express **26**, 17360–17377 (2018).

S. Civelli, J. E. Prilepsky, M. Secondini, and S. K. Turitsyn, "Polarization division multiplexing for nonlinear Fourier-based transmission schemes", in "CLEO: Science and Innovations," (Optical Society of America, 2018), STu4C–4 (2018).

S. Civelli, E. Forestieri, and M. Secondini, "Decision-feedback detection strategy for nonlinear frequency-division multiplexing", Optics Express **26**, 12057–12071 (2018).

S. Civelli, E. Forestieri, and M. Secondini, "A novel detection strategy for nonlinear frequency-division multiplexing," in "Optical Fiber Communication Conference," (Optical Society of America, 2018), W1G–5 (2018).

S. Civelli, E. Forestieri, and M. Secondini, "Why noise and dispersion may seriously hamper nonlinear frequency-division multiplexing," IEEE Photonics Technology Letters **29**, 1332–1335 (2017).

S. Civelli, E. Forestieri, and M. Secondini, "Precompensation and windowing for nonlinear frequency-division multiplexing," in "Progress In Electromagnetics Research Symposium - Spring (PIERS) 2017," (2017).

S. Civelli, E. Forestieri, and M. Secondini, "Impact of discretizations and boundary conditions in nonlinear frequency-domain multiplexing," in "18th Italian National Conference on Photonic Technologies (Fotonica 2016)," (2016).

S. Civelli, L. Barletti, and M. Secondini, "Numerical methods for the inverse nonlinear Fourier transform," in "Proc. Tyrrhenian Int. Workshop Digital Communications (TIWDC)," (Florence, Italy, 2015), pp. 13–16 (2015).

Bibliography

[1] G. P. Agrawal, *Fiber-optic communication systems* (John Wiley & Sons, 2002).

[2] K. Kao and G. A. Hockham, "Dielectric-fibre surface waveguides for optical frequencies," in "Proceedings of the Institution of Electrical Engineers," , vol. 113 (IET, 1966), vol. 113, pp. 1151–1158.

[3] E. Agrell, M. Karlsson, A. Chraplyvy, D. J. Richardson, P. M. Krummrich, P. Winzer, K. Roberts, J. K. Fischer, S. J. Savory, B. J. Eggleton *et al.*, "Roadmap of optical communications," Journal of Optics **18**, 063002 (2016).

[4] D. J. Richardson, "Filling the light pipe," Science **330**, 327–328 (2010).

[5] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," Journal of Lightwave Technology **28**, 662–701 (2010).

[6] J. C. Cartledge, F. P. Guiomar, F. R. Kschischang, G. Liga, and M. P. Yankov, "Digital signal processing for fiber nonlinearities," Optics Express **25**, 1916–1936 (2017).

[7] S. K. Turitsyn, J. E. Prilepsky, S. T. Le, S. Wahls, L. L. Frumin, M. Kamalian, and S. A. Derevyanko, "Nonlinear Fourier transform for optical data processing and transmission: advances and perspectives," Optica **4**, 307–322 (2017).

[8] M. J. Ablowitz and H. Segur, *Solitons and the inverse scattering transform*, vol. 4 (SIAM, 1981).

[9] M. J. Ablowitz, B. Prinari, and A. D. Trubatch, *Discrete and Continuous Nonlinear Schroedinger Systems* (Cambridge University Pres, 2004).

[10] A. Shabat and V. Zakharov, "Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media," Soviet physics-Journal of Experimental and Theoretical Physics (JETP) **34**, 62 (1972).

[11] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear Fourier transform, Parts I–III," IEEE Transactions on Information Theory **60**, 4312–4369 (2014).

[12] M. I. Yousefi and X. Yangzhang, "Linear and nonlinear frequency-division multiplexing," European Conference on Optical Communication 2016 (ECOC) (2016).

[13] H. Bülow, "Experimental demonstration of optical signal detection using nonlinear Fourier transform," Journal of Lightwave Technology **33**, 1433–1439 (2015).

[14] V. Aref, S. T. Le, and H. Buelow, "Modulation over nonlinear Fourier spectrum: Continuous and discrete spectrum," Journal of Lightwave Technology **36**, 1289–1295 (2018).

[15] I. Tavakkolnia and M. Safari, "Capacity analysis of signaling on the continuous spectrum of nonlinear optical fibers," Journal of Lightwave Technology **35**, 2086–2097 (2017).

[16] S. T. Le, I. D. Philips, J. E. Prilepsky, P. Harper, A. D. Ellis, and S. K. Turitsyn, "Demonstration of nonlinear inverse synthesis transmission over transoceanic distances," Journal of Lightwave Technology **34**, 2459–2466 (2016).

[17] M. Kamalian, J. E. Prilepsky, S. T. Le, and S. K. Turitsyn, "Periodic nonlinear Fourier transform for fiber-optic communications, Part I-II," Optics express **24**, 18353–18381 (2016).

[18] S. T. Le, V. Aref, and H. Buelow, "Nonlinear signal multiplexing for communication beyond the Kerr nonlinearity limit," Nature Photonics **11**, 570 (2017).

[19] J. E. Prilepsky, S. A. Derevyanko, K. J. Blow, I. Gabitov, and S. K. Turitsyn, "Nonlinear inverse synthesis and eigenvalue division multiplexing in optical fiber channels," Physical review letters **113**, 013901 (2014).

[20] L. Mollenauer and K. Smith, "Demonstration of soliton transmission over more than 4000 km in fiber with loss periodically compensated by Raman gain," Optics letters **13**, 675–677 (1988).

[21] M. Nakazawa, E. Yamada, H. Kubota, and K. Suzuki, "10 Gbit/s soliton data transmission over one million kilometres," Electronics Letters **27**, 1270–1272 (1991).

[22] A. Hasegawa and T. Nyu, "Eigenvalue communication," Journal of Lightwave Technology **11**, 395–399 (1993).

[23] E. V. Sedov, A. A. Redyuk, M. P. Fedoruk, A. A. Gelash, L. L. Frumin, and S. K. Turitsyn, "Soliton content in the standard optical ofdm signal," Optics Letters **43**, 5985–5988 (2018).

[24] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, "Korteweg-deVries equation and generalizations. VI. methods for exact solution," Communications on pure and applied mathematics **27**, 97–133 (1974).

[25] "2006 Steele Prizes," Notices of the AMS **53**, 464–470 (2006).

[26] F. Da Ros, S. Civelli, S. Gaiarin, E. P. da Silva, N. De Renzis, M. Secondini, and D. Zibar, "Dual-polarization NFDM transmission with continuous and discrete spectral modulation," Journal of Lightwave Technology (2019).

[27] C. E. Shannon, "A mathematical theory of communication," Bell System Technical Journal **27**, 623–656 (1948).

[28] T. M. Cover and J. A. Thomas, *Elements of information theory* (John Wiley & Sons, 2012).

[29] G. P. Agrawal, *Nonlinear Fiber Optics* (Academic Press, San Diego, CA, 2001), 3rd ed.

[30] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn, "Nonlinear inverse synthesis technique for optical links with lumped amplification," Optics express **23**, 8317–8328 (2015).

[31] C. R. Menyuk and B. S. Marks, "Interaction of polarization mode dispersion and nonlinearity in optical fiber transmission systems," Journal of Lightwave Technology **24**, 2806 (2006).

[32] C. Menyuk, "Application of multiple-length-scale methods to the study of optical fiber transmission," Journal of Engineering Mathematics **36**, 113–136 (1999).

[33] S. V. Manakov, "On the theory of two-dimensional stationary self-focusing of electromagnetic waves," Soviet Physics-Journal of Experimental and Theoretical Physics (JETP) **38**, 248–253 (1974).

[34] A. Mecozzi, C. Antonelli, and M. Shtaif, "Nonlinear propagation in multimode fibers in the strong coupling regime," Optics express **20**, 11673–11678 (2012).

[35] S. Mumtaz, R.-J. Essiambre, and G. P. Agrawal, "Nonlinear propagation in multimode and multicore fibers: generalization of the Manakov equations," Journal of Lightwave Technology **31**, 398–406 (2013).

[36] L. B. Du, M. M. Morshed, and A. J. Lowery, "Fiber nonlinearity compensation for OFDM super-channels using optical phase conjugation," Optics express **20**, 19921–19927 (2012).

[37] E. Grellier and A. Bononi, "Quality parameter for coherent transmissions with Gaussian-distributed nonlinear noise," Optics Express **19**, 12781–12788 (2011).

[38] R. A. Shafik, M. S. Rahman, and A. R. Islam, "On the extended relationships among EVM, BER and SNR as performance metrics," in "International Conference on Electrical and Computer Engineering, 2006 (ICECE'06)," (IEEE, 2006), pp. 408–411.

[39] J. G. Proakis and M. Salehi, *Digital Communications (5th ed.)* (McGraw-Hill, 2008).

[40] S. Civelli, S. K. Turitsyn, M. Secondini, and J. E. Prilepsky, "Polarization-multiplexed nonlinear inverse synthesis with standard and reduced-complexity NFT processing," Optics Express **26**, 17360–17377 (2018).

[41] S. Civelli, L. Barletti, and M. Secondini, "Numerical methods for the inverse nonlinear Fourier transform," in "2015 Tyrrhenian International Workshop on Digital Communications (TIWDC)," (IEEE, 2015), pp. 13–16.

[42] J. S. Russel, "Report of the committee on waves," in "Report of the 7th Meeting of British Association for the Advancement of Science," (1838), pp. 417–496.

[43] D. J. Korteweg and G. De Vries, "XLI. on the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves," The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **39**, 422–443 (1895).

[44] N. J. Zabusky and M. D. Kruskal, "Interaction of" solitons" in a collisionless plasma and the recurrence of initial states," Physical Review Letters **15**, 240 (1965).

[45] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, "Method for solving the Korteweg-deVries equation," Physical Review Letters **19**, 1095 (1967).

[46] P. D. Lax, "Integrals of nonlinear equations of evolution and solitary waves," Communications on pure and applied mathematics **21**, 467–490 (1968).

[47] T. Aktosun, "Inverse scattering transform and the theory of solitons," in "Encyclopedia of Complexity and Systems Science," (Springer, 2009), pp. 4960–4971.

[48] F. Demontis, "Direct and inverse scattering of the matrix Zakharov-Shabat system," Ph.D. thesis, Università degli studi di Cagliari (2006).

[49] J.-W. Goossens, M. I. Yousefi, Y. Jaouën, and H. Hafermann, "Polarization-division multiplexing based on the nonlinear fourier transform," Optics express **25**, 26437–26452 (2017).

[50] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn, "Nonlinear inverse synthesis for high spectral efficiency transmission in optical fibers," Optics Express **22**, 26720–26741 (2014).

[51] F. J. G. Gomez, "Communication using eigenvalues of higher multiplicity of the nonlinear fourier transform," Journal of Lightwave Technology (2018).

[52] M. Kamalian, A. Vasylchenkova, D. Shepelsky, J. E. Prilepsky, and S. K. Turitsyn, "Signal modulation and processing in nonlinear fibre channels by employing the Riemann-Hilbert problem," Journal of Lightwave Technology (2018).

[53] O. Wright, "The Darboux transformation of some Manakov systems," Applied mathematics letters **16**, 647–652 (2003).

[54] S. Civelli, E. Forestieri, and M. Secondini, "Decision-feedback detection strategy for nonlinear frequency-division multiplexing," Optics Express **26**, 12057–12071 (2018).

[55] O. Belai, L. Frumin, E. Podivilov, and D. Shapiro, "Efficient numerical method of the fiber Bragg grating synthesis," JOSA B **24**, 1451–1457 (2007).

[56] I. T. Lima, T. D. DeMenezes, V. Grigoryan, M. O'sullivan, and C. R. Menyuk, "Nonlinear compensation in optical communications systems with normal dispersion fibers using the nonlinear Fourier transform," Journal of Lightwave Technology (2017).

[57] A. Rosenthal and M. Horowitz, "Inverse scattering algorithm for reconstructing strongly reflecting fiber Bragg gratings," IEEE Journal of quantum electronics **39**, 1018–1026 (2003).

[58] G. Boffetta and A. R. Osborne, "Computation of the direct scattering transform for the nonlinear Schroedinger equation," Journal of Computational Physics **102**, 252–264 (1992).

[59] S. Wahls and H. V. Poor, "Fast numerical nonlinear Fourier transforms," IEEE Transactions on Information Theory **61**, 6957–6974 (2015).

[60] S. Wahls, S. Chimmalgi, and P. J. Prins, *FNFT: A Software Library for Computing Nonlinear Fourier Transforms.*

[61] E. Bidaki and S. Kumar, "Nonlinear Fourier transform using multistage perturbation technique for fiber-optic systems," Journal of the Optical Society of America (JOSA) B **35**, 2286–2293 (2018).

[62] A. Vasylchenkova, J. E. Prilepsky, and S. K. Turitsyn, "Contour integrals for numerical computation of discrete eigenvalues in the Zakharov–Shabat problem," Optics letters **43**, 3690–3693 (2018).

[63] A. Vasylchenkova, J. Prilepsky, D. Shepelsky, and A. Chattopadhyay, "Direct nonlinear Fourier transform algorithms for the computation of solitonic spectra in focusing nonlinear Schrödinger equation," Communications in Nonlinear Science and Numerical Simulation **68**, 347–371 (2019).

[64] S. A. Derevyanko, J. E. Prilepsky, and S. K. Turitsyn, "Capacity estimates for optical transmission based on the nonlinear Fourier transform," Nature Communications **7**, 12710 (2016).

[65] A. Aricò, G. Rodriguez, and S. Seatzu, "Numerical solution of the nonlinear Schrödinger equation, starting from the scattering data," Calcolo **48**, 75–88 (2011).

[66] L. L. Frumin, O. V. Belai, E. V. Podivilov, and D. A. Shapiro, "Efficient numerical method for solving the direct Zakharov–Shabat scattering problem," Journal of the Optical Society of America (JOSA) B **32**, 290–296 (2015).

[67] V. Aref, "Control and detection of discrete spectral amplitudes in nonlinear Fourier spectrum," arXiv preprint arXiv:1605.06328 (2016).

[68] S. Gaiarin, A. M. Perego, E. P. da Silva, F. Da Ros, and D. Zibar, "Dual-polarization nonlinear Fourier transform-based optical communication system," Optica **5**, 263–270 (2018).

[69] S. Civelli, S. Turitsyn, M. Secondini, and J. Prilepsky, "Polarization-multiplexed nonlinear inverse synthesis with standard and reduced-complexity NFT processing: erratum," Optics express **27**, 3617–3617 (2019).

[70] S. Gaiarin, A. M. Perego, E. P. da Silva, F. Da Ros, and D. Zibar, "Experimental demonstration of dual polarization nonlinear frequency division multiplexed optical transmission system," in "European Conference on Optical Communication 2017 (ECOC)," (2017), p. W.3.C.2.

[71] A. Maruta and Y. Matsuda, "Polarization division multiplexed optical eigenvalue modulation," in "2015 International Conference on Photonics in Switching (PS)," (IEEE, 2015), pp. 265–267.

[72] T. Gui, T. H. Chan, C. Lu, A. P. T. Lau, and P.-K. A. Wai, "Alternative decoding methods for optical communications based on nonlinear fourier transform," Journal of Lightwave Technology **35**, 1542–1550 (2017).

[73] T. Gui, G. Zhou, C. Lu, A. P. T. Lau, and S. Wahls, "Nonlinear frequency division multiplexing with b-modulation: shifting the energy barrier," Optics Express **26**, 27978–27990 (2018).

[74] A. Hause, C. Mahnke, and F. Mitschke, "Impact of fiber loss on two-soliton states: Substantial changes in eigenvalue spectrum," Physical Review A **98**, 033814 (2018).

[75] S. Civelli, E. Forestieri, and M. Secondini, "Why noise and dispersion may seriously hamper nonlinear frequency-division multiplexing," IEEE Photonics Technology Letters **29**, 1332–1335 (2017).

[76] S. Civelli, E. Forestieri, and M. Secondini, "A novel detection strategy for nonlinear frequency-division multiplexing," in "Optical Fiber Communication Conference," (Optical Society of America, 2018), pp. W1G–5.

[77] S. Civelli, E. Forestieri, and M. Secondini, "Improved detection strategies for nonlinear frequency-division multiplexing," arXiv preprint arXiv:1805.02047, Progress In Electromagnetics Research Symposium (PIERS) 2018 (2018).

[78] S. Civelli, E. Forestieri, and M. Secondini, "Impact of discretizations and boundary conditions in nonlinear frequency-domain multiplexing," in "18th Italian National Conference on Photonic Technologies (Fotonica 2016)," (2016).

[79] S. Civelli, E. Forestieri, and M. Secondini, "Precompensation and windowing for nonlinear frequency-division multiplexing," in "Progress In Electromagnetics Research Symposium - Spring (PIERS) 2017," (2017).

[80] H. Buelow, V. Aref, and W. Idler, "Transmission of waveforms determined by 7 eigenvalues with PSK-modulated spectral amplitudes," in "European Conference on Optical Communication 2016 (ECOC)," (VDE, 2016), pp. 1–3.

[81] S. Li, J. Koch, and S. Pachnicke, "Optical signal processing in the discrete nonlinear frequency domain," in "Optical Fiber Communication Conference," (Optical Society of America, 2018), pp. 1–3.

[82] Z. Dong, S. Hari, T. Gui, K. Zhong, M. Yousefi, C. Lu, P.-K. A. Wai, F. Kschischang, and A. Lau, "Nonlinear frequency division multiplexed transmissions based on NFT," IEEE Photonics Technology Letters **99**, 1 (2015).

[83] T. Gui, C. Lu, A. P. T. Lau, and P. Wai, "High-order modulation on a single discrete eigenvalue for optical communications based on nonlinear fourier transform," Optics express **25**, 20286–20297 (2017).

[84] V. Aref, S. T. Le, and H. Buelow, "Demonstration of fully nonlinear spectrum modulated system in the highly nonlinear optical transmission regime," European Conference on Optical Communication 2016 (ECOC) (2016).

[85] I. Tavakkolnia and M. Safari, "Dispersion pre-compensation for NFT-based optical fiber communication systems," in "Conference on Lasers and Electro-Optics (CLEO)," (Optical Society of America, 2016).

[86] R. T. Jones, S. Gaiarin, M. P. Yankov, and D. Zibar, "Noise robust receiver for eigenvalue communication systems," in "Optical Fiber Communication Conference," (Optical Society of America, 2018).

[87] T. Gui, W. A. Gemechu, J.-W. Goossens, M. Song, S. Wabnitz, M. I. Yousefi, H. Hafermann, A. P. T. Lau, and Y. Jaouën, "Polarization-division-multiplexed nonlinear frequency division multiplexing," in "Conference on Lasers and Electro-Optics (CLEO)," (Optical Society of America, 2018), pp. STu4C–3.

[88] S. Gaiarin, F. Da Ros, N. De Renzis, E. P. da Silva, and D. Zibar, "Dual-polarization nfdm transmission using distributed raman amplification and nft-domain equalization," IEEE Photonics Technology Letters (2018).

[89] S. Wahls, "Generation of time-limited signals in the nonlinear Fourier domain via b-modulation," in "European Conference on Optical Communication 2017 (ECOC)," (IEEE, 2017), pp. 1–3.

[90] S. T. Le, K. Schuh, F. Buchali, and H. Bülow, "100 gbps b-modulated nonlinear frequency division multiplexed transmission," in "Optical Fiber Communication Conference," (Optical Society of America, 2018), pp. W1G–6.

[91] X. Yangzhang, V. Aref, S. T. Le, H. Bülow, and P. Bayvel, "400 Gbps dual-polarisation non-linear frequency-division multiplexed transmission with b-modulation," in "2018 European Conference on Optical Communication (ECOC)," (IEEE, 2018), pp. 1–3.

[92] L. Faddeev and L. Takhtajan, *Hamiltonian methods in the theory of solitons* (Springer Science & Business Media, 2007).

[93] V. Aref, S. T. Le, and H. Buelow, "An efficient nonlinear Fourier transform algorithm for detection of eigenvalues from continuous spectrum," arXiv preprint arXiv:1812.02092 (2018).

[94] N. A. Shevchenko, S. A. Derevyanko, J. E. Prilepsky, A. Alvarado, P. Bayvel, and S. K. Turitsyn, "Capacity lower bounds of the noncentral chi-channel with applications to soliton amplitude modulation," IEEE Transactions on Communications **66** (2018).

[95] S. Civelli, J. E. Prilepsky, M. Secondini, and S. K. Turitsyn, "Polarization division multiplexing for nonlinear Fourier-based transmission schemes," in "Conference on Lasers and Electro-Optics (CLEO)," (Optical Society of America, 2018), pp. STu4C–4.

[96] S. Civelli, S. K. Turitsyn, M. Secondini, and J. E. Prilepsky, "Reduced complexity nonlinear inverse synthesis transmission method with polarization division multiplexing," in "European Conference on Optical Communication 2018 (ECOC)," (2018).

[97] W. A. Gemechu, T. Gui, J.-W. Goossens, M. Song, S. Wabnitz, H. Hafermann, A. P. T. Lau, M. I. Yousefi, and Y. Jaouën, "Dual polarization nonlinear frequency division multiplexing transmission," IEEE Photonics Technology Letters **30**, 1589–1592 (2018).

[98] A. Geisler, J. Leibrich, and C. G. Schaeffer, "Influence of non-ideal first order counter-propagating Raman amplification on discrete nonlinear Fourier spectrum based communication," in "Photonic Networks; 19th ITG-Symposium," (VDE, 2018), pp. 1–7.

[99] V. Aref, H. Bülow, K. Schuh, and W. Idler, "Experimental demonstration of nonlinear frequency division multiplexed transmission," in "European Conference on Optical Communication 2015 (ECOC)," (IEEE, 2015), pp. 1–3.

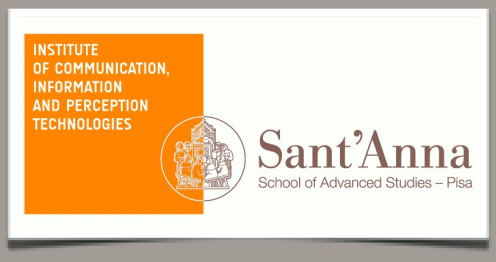
[100] A. Span, V. Aref, H. Bülow, and S. Ten Brink, "On time-bandwidth product of multi-soliton pulses," in "Information Theory (ISIT), 2017 IEEE International Symposium on," (IEEE, 2017), pp. 61–65.

[101] E. Agrell and M. Secondini, "Information-theoretic tools for optical communications engineers," in "Proceedings of IEEE Photonic Conference 2018 (IPC 2018)," (2018), p. MA3.1.

[102] H. Sun, K.-T. Wu, and K. Roberts, "Real-time measurements of a 40 gb/s coherent system," Optics Express **16**, 873–879 (2008).

[103] V. Aref, S. T. Le, and H. Buelow, "Does the cross-talk between nonlinear modes limit the performance of nfdm systems?" in "European Conference on Optical Communication 2017 (ECOC)," (2017).

[104] A. Span, V. Aref, H. Buelow, and S. t. Brink, "Precoding for dual polarization soliton transmission," arXiv preprint arXiv:1808.09904 (2018).


INSTITUTE
OF COMMUNICATION,
INFORMATION
AND PERCEPTION
TECHNOLOGIES
Sant'Anna
School of Advanced Studies – Pisa