Technische Universität München
TUM School of Computation, Information and Technology

Fiber-Optic Communication via the Nonlinear Fourier Transform

Benedikt Leible

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung eines

Doktors der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Gabriele Schrag

Prüfende der Dissertation:

1. Prof. Dr.-Ing. Norbert Hanik
2. Prof. Dr.-Ing. Stephan ten Brink

Die Dissertation wurde am 29.04.2024 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 31.10.2024 angenommen.

Fiber-Optic Communication via the Nonlinear Fourier Transform

Leible, Benedikt

December 29, 2024

Contents

1. Introduction **9**

1.1. Capacity Crunch in Fiber-Optic Systems 10

1.2. Solitons, Inverse Scattering Theory and NFDM-Transmission 11

1.3. Challenges of NFDM System-Design and Focus of this Thesis 12

2. Preliminaries **15**

2.1. Definitions and Notation 15

2.1.1. Special Sets and Subsets 15

2.1.2. Vectors, Matrices and Operators 16

2.1.3. Subscripts and Superscripts 16

2.1.4. Functions and Transformations 17

2.1.5. Random Variables 18

2.2. Aspects of Information Theory 18

2.2.1. Basic Metrics 19

2.2.2. Upper Bounds on Mutual Information 20

2.3. Aspects of Operator Theory 20

2.3.1. The Adjoint and Inverse Operator 21

2.3.2. The Spectrum of an Operator 21

2.4. Aspects of Perturbation Theory 22

2.5. Performance Metrics 23

2.5.1. Mean Squared Error 23

2.5.2. Bit Error-Rate 24

2.5.3. Mismatched Mutual Information 25

2.5.4. Achievable Information Rate and Spectral Efficiency 25

3. Fiber-Optic Communication Systems **27**
3.1. The Single-Mode Fiber Channel 28
3.1.1. Fiber-Loss 29
3.1.2. Dispersion 30
3.1.3. Kerr-Nonlinearity 31
3.1.4. Polarization Dependent Effects 33
3.1.5. Numerical Solution by Split-Step Analysis 35
3.2. The Strong Coupling Multi-Mode Fiber Channel 39
3.3. Amplification Schemes 40
3.3.1. Lumped Amplification by Erbium Doped Fiber Amplifiers . . . 40
3.3.2. Distributed Amplification by Raman Amplifiers 42
3.4. Modulation Formats and Bit-Labeling 47
3.4.1. QAM Constellation 47
3.4.2. MPSK Constellation 48
3.4.3. Bit-Labeling Schemes 49
3.5. Wave-Division Multiplexing Transmission Systems 51
3.5.1. Transmission and Detection 52
3.5.2. Digital Signal Processing 54
3.5.3. Lower Bounds on Achievable Information Rates 55

4. The Nonlinear Fourier Transform for Single-Mode Fiber Channels **59**
4.1. Basic Concept and Transformations 60
4.1.1. Lax Pairs for Nonlinear Evolution Equations 61
4.1.2. The Zero-Curvature Condition 62
4.1.3. The Zakharov-Shabat System 63
4.1.4. Boundary Conditions and the Wronskian 63
4.1.5. Nonlinear Fourier Coefficients in the Complex Plane 65
4.1.6. The Nonlinear Fourier Spectra 67
4.1.7. The Lax Convolution Channel 68
4.1.8. The Nonlinear Fourier Transform 68
4.2. Properties of the Nonlinear Spectra 69
4.2.1. Phase, Time and Frequency Shifts 69
4.2.2. The Layer-Peeling Property 70
4.2.3. The Trace Formula 71
4.2.4. Propagation of Nonlinear Spectra in the NLSE Channel 72

4.3. Special Solutions . . . 73
4.3.1. The Rectangular Pulse . . . 73
4.3.2. The Fundamental Soliton . . . 75
4.3.3. The Satsuma-Yajima Pulse . . . 79
4.4. Algorithms . . . 80
4.4.1. Forward NFT for the Continuous Spectrum . . . 80
4.4.2. Forward NFT for the Discrete Spectrum . . . 86
4.4.3. The Eigenvalue Removal NFT . . . 98
4.4.4. Coefficient Recovery from Nonlinear Spectra . . . 107
4.4.5. Inverse NFT for the Continuous Spectrum . . . 109
4.4.6. Inverse NFT for the Discrete Spectrum . . . 112
4.4.7. Full Nonlinear Spectrum INFT . . . 114
4.5. On the Time-Bandwidth Product of Solitons . . . 116
4.5.1. TBR for Multi-Ring Modulated Trains of Fundamental Solitons 117
4.5.2. TBR for Multi-Ring Modulated Trains of Multi-Solitons . . . 119
4.6. Improved Constellations for Timing-Jitter Resistant Soliton Trains . . . 123
4.7. Eigenvalue Trajectories in Lossy Channels . . . 125
4.7.1. Trajectory Under Lumped Amplification . . . 126
4.7.2. Trajectory Under Distributed Amplification . . . 127
4.8. Collision and Absorption Phenomena . . . 128
4.9. Path-Loss Averaged Transmission . . . 129
4.9.1. PLA-Enhanced Transmission over EDFA Channels . . . 130
4.9.2. PLA-Enhanced Transmission over DRA Channels . . . 131
4.10. Fiber Parameter Estimation using the NFT . . . 133
4.10.1. Nonlinearity Estimation with Fundamental Solitons . . . 133
4.10.2. Joint Parameter Estimation with Multi-Solitons . . . 135
4.11. Labeling Approaches for EV-OOK Systems . . . 139
4.11.1. 'Gray' Labeling . . . 141
4.11.2. Greedy Labeling . . . 141
4.11.3. Simulation Results . . . 142
4.12. Clustering Algorithm for Discrete Spectrum Data . . . 143
4.13. Soliton Receivers with Matched-Filters . . . 147

5. The Nonlinear Fourier Transform for Multi-Mode Fibers 151
5.1. Basic Concept and Transformations . . . 152

5.2. Properties of the Nonlinear Spectra . . . 154
5.2.1. Constant Phase Change . . . 154
5.2.2. Time Shift . . . 155
5.2.3. Trace Formula . . . 155
5.2.4. Layer-Peeling Property . . . 156
5.2.5. Propagation of Nonlinear Quantities . . . 157
5.3. Special Solutions . . . 157
5.3.1. The Rectangular Pulse-Set . . . 157
5.3.2. The Fundamental Soliton . . . 158
5.4. Algorithms . . . 161
5.4.1. Forward Transformations for the Continuous Spectrum . . . 162
5.4.2. Forward Transformations for the Discrete Spectrum . . . 164
5.4.3. Inverse NFT for the Discrete Spectrum . . . 167

6. Conclusions **171**
6.1. Overview . . . 171
6.2. Future Work . . . 173
6.3. Acknowledgement . . . 175

A. Additional Proofs **177**
A.1. The Lax Equation . . . 177
A.2. The Zero Curvature Condition . . . 178
A.3. Properties of the Fundamental Soliton . . . 179
A.3.1. Temporal Width . . . 179
A.3.2. Bandwidth . . . 179
A.4. 'Gray' Labeling . . . 180
A.5. Properties of Nonlinear Spectra in the Multi-Mode Case . . . 182
A.5.1. Differential Representation for Multi-Mode NFT . . . 182
A.5.2. Constant Phase Change . . . 185
A.5.3. Time Shift . . . 185
A.5.4. Trace Formula . . . 186
A.6. NFT of the Vectorial Rectangular Pulse . . . 187
A.7. Properties of the Vectorial Fundamental Soliton . . . 189
A.7.1. Temporal Width . . . 189
A.7.2. Bandwidth . . . 190

A.8. Algorithms for the Multi-Mode Case 191

B. Additional Algorithms **193**
B.1. Alternative Continuous Spectrum RH-Solver 193
B.2. Greedy Labeling Algorithm 194
B.3. Marching Squares Algorithm 195
B.4. Fiber Parameter Estimation Algorithm 196

C. Additional Parameter Tables **199**
C.1. Table for WDM Simulations 199
C.2. Table for PLA-Enhanced Transmission Systems 200
C.3. Table for Fiber Parameter Estimation with Multiple Discrete Eigenvalues 200
C.4. Table for EM-Clustering Results 201

D. Additional Simulation Results **203**
D.1. Additional Simulation Results for WMD Systems 203
D.2. Accuracy of the Continuous Spectrum NFT 204
D.3. Accuracy of the Eigenvalue Removal NFT 206
D.4. PLA-Enhanced Multi-Soliton Systems 209
D.5. Fiber Parameter Estimation 210
D.6. EV-OOK Labeling Schemes 212
D.7. Clustering Algorithm for Discrete Spectrum Data 213
D.7.1. Example for Clustering with the EM Algorithm 214
D.7.2. Results for K-Means Clustering 215

E. Notation, Common Functions, Symbols and Acronyms **217**
E.1. Notation and Functions 217
E.2. Symbols 219
E.3. Acronyms 225

F. List of Publications **231**

Bibliography **233**

1 Introduction

Exchanging information, in spoken or written form, in a singularly nuanced way can be seen as an essential key feature that enabled humanity to progress to the high technological level it currently is at. To facilitate the fast exchange of information over long distances, humanity started to invent means of *telecommunication* very early in history, reducing the need for the exchange of information by travel. The first construct that probably could be called an optical long-distance communication system was used by the Greek army to signal the capture of Troy in 1180BC over a distance of more than 600km [1]. Fueled by the advances in optics for telescopes in the late 18th century, Claude Chappe invented a system for optical communication, utilizing an elaborate apparatus consisting of movable wooden beams on top of buildings in elevated positions. He developed a system with an elaborate transmission alphabet of 98 different symbols. After testing the first transmission line from Lille to Paris, this led to a countrywide optical telegraphy network consisting of 556 telegraph stations, spanning a total distance of approximately 4800 km and linking 29 of France's largest cities with Paris [2]. During this endeavor, the term *telegraphy* was coined for the first time.

Compared to these early efforts, data transmission over fiber-optic links is a very recent development that gained momentum with the invention of the laser, the low-loss single mode fiber (SMF), and the erbium doped fiber amplifier (EDFA) between the years 1960 and 1990 [3]. Even though coherent receivers were extensively studied in the period between 1980 and 1992, they first gained commercial recognition around

2005 [4]. Coherent receivers are the key technology enabling higher-order modulation schemes that utilize phase information in addition to the already used magnitude of the signal. With the low-loss SMF, increasing the reach of fiber-optic systems, and the EDFA, enabling fully optical transparent signal regeneration, system reach as well as data rates increased significantly over a very short period of time.

By 2003, a global network of transatlantic and transpacific fiber-optic transmission lines was built, encompassing a total of over 600 000km of fiber, connecting the continents [5]. In 1998, the volume of data traffic exceeded the volume of voice traffic for the first time. Since then, the transmission of data has been a key driver for the increase in maximum data rates in fiber-optic systems.

1.1. Capacity Crunch in Fiber-Optic Systems

Even though, data-rates for fiber-optic systems initially increased strongly, due to the utilization of EDFAs in conjunction with low-loss fibers and coherent detection, the achievable rates began to saturate in recent years. In contrast, the traffic growth is projected to be 30-60% per year [6]. Currently achievable spectral efficiencys (SEs) vary depending on the system for which they are measured. An overview of recent results in this field for dual-polarization and space division multiplexing (SDM) systems can be found in [7], [8] and [9] respectively. They show that e.g. a SE of ≈ 9 bit/s/Hz/polarization is achievable for a system with a reach of $\mathcal{L} = 1000$km assuming ideal Raman amplification. Using lumped amplification with a span length of 60km for a total transmission distance of $\mathcal{L} = 1000$km and $\mathcal{L} = 4000$km yields a maximum SE of ≈ 7 and ≈ 5.5 bit/s/Hz/polarization respectively [7]. Still, with the growth rates predicted in [6], [10] it is clear to see, that there is a mismatch between the growth of the required and available data rates. This phenomenon is commonly known as the fiber-optic networks *capacity crunch*.

One widely studied approach to combat this phenomenon is extending the utilized frequency band. While this does generally not increase the SE of a system, the data-rates can be scaled by using more of the available frequency spectrum for transmission. Instead of only using the C-band, including the neighboring S- and L-band and potentially even adding the O- and E-band using dry-fibers is considered [4], [11]. Another widely studied idea is scaling the data-rates by using multiple spatial paths. Apart from simply using more fiber links in parallel, this also includes the study of multi-

mode and multi-core fibers. Those two approaches can also be used simultaneously, to achieve even better scaling of data-rates in fiber-optic transmission systems [4], [12]. Both of these approaches require exchanging or extending parts of the already existing fiber-optic network.

Another approach, is the use of the nonlinear Fourier transform (NFT), to modulate decoupled linear modes of transmission in the nonlinear Fourier domain (NFD), instead of utilizing the linear frequency domain, in which channels interact during propagation, due to fiber nonlinearity. In many cases, this approach does not require significant physical changes to the deployed network structure. The NFT-based modulation and detection schemes can be added by primarily making alterations to the digital signal processing (DSP) sections of a transmitter/receiver pair.

While SE gains from modulation in the NFD, are limited by the upper bound derived in [13], the margin for improvement due to the use of the NFT is rather high. This is especially true for communication in the highly-nonlinear power regime of systems, for which the achievable information rates tend to drop off for state-of-the-art wave division multiplexing (WDM) systems due to nonlinear signal-noise interactions even if advanced DSP is used at the receiver. Further, the concept of NFT-aided transmission can generally be combined with approaches like SDM and thus does not have any immediate drawbacks in terms of compatibility with other approaches mentioned in this section.

1.2. Solitons, Inverse Scattering Theory and NFDM-Transmission

In 1834, a phenomenon linked to the NFT was documented for the first time, when John Scott Russell observed a water wave, that propagated along a canal in a steady pace without significantly changing its shape for several kilometers [14]. He later also observed, that upon other remarkable properties, two such waves seemed to propagate through each other without significantly changing their shape after collision. This particle-like behavior lead to Russell calling them solitary waves. This *solitons*, as they were later called, are also observable in optical fibers and were predicted by *inverse scattering theory*, which can be applied to many different problems in physics and mathematics. In 1971, Zakharov and Shabat successfully applied this theory to the nonlinear Schrödinger equation (NLSE), which governs the propagation of light

through optical SMF [15]. Two years later, the first simulations showing the stability of solitons were presented by [16], [17] and with the publication of Lax [18], the theory of inverse scattering and soliton transmission was solidified mathematically. The name NFT was later introduced as a nod to the parallels between the linear Fourier transform in the context of linear channels and the NFT in the context of channels that exhibit a certain nonlinear behavior, such as the NLSE.

Since then, a plethora of publications have been written, studying the NFT in regards to fiber-optic communication. This includes papers that deal exclusively with soliton transmission and thus the discrete nonlinear Fourier spectrum [19]–[22], the continuous or radiative nonlinear Fourier spectrum [23], [24], experimental evaluations [25], [26] and advanced algorithms, as well as reviews of known numerical techniques [27]–[31].

1.3. Challenges of NFDM System-Design and Focus of this Thesis

The underlying idea of signal-modulation in a domain in which, in theory, the different nonlinear carrier-frequencies do not interact which each other, even though being rather elegant mathematically, also generates some new problems diminishing the anticipated performance of such systems [32], [33]. This includes the mitigation of effects, neglected in the underlying theory, such as fiber-loss, distributed or lumped noise, polarization mode dispersion (PMD), and hardware impairments in transmitter and receiver parts. Another challenge, is making NFT-aided transmission schemes compatible with technologies that potentially will be widely deployed in the future, such as e.g. SDM.

This thesis is structured as follows: After this introduction, a short review of some necessary preliminaries is given in chapter 2. This includes an overview of the used definitions and metrics as well as some conventions regarding notation. Further, some aspects from the fields of information, operator and perturbation theory, that are utilized in later sections, are reviewed shortly. In the next chapter, an overview of common building blocks of fiber-optic transmission systems is given and WDM systems in the point-to-point and network scenario are discussed. Chapter 3 also includes some results on the achievable information rates of such systems. In chapter 4, the nonlinear Fourier transform for the single-polarization lossless Schrödinger equation is

studied. This includes a review of the mathematical basis of the forward and backward transformations, the discussion of several provable properties of the nonlinear Fourier transformation, some results for special pulse-shapes and a large section on algorithms for forward and inverse transformations. After that, some more in-depth topics regarding the single-polarization case are discussed. This includes e.g. a treatment of fiber parameter estimation algorithms utilizing the nonlinear Fourier transform, as well as some approaches on improving the transmission quality of NFT-aided systems by using labeling and clustering algorithms at the receiver. In chapter 5, many of the obtained results from chapter 4 are extended to the strong coupling multi-mode case. This includes the derivation of the necessary mathematical basis for the transformations linked to this specific channel model, as well as basic properties and results for special pulse-sets. Also several algorithms, which are well known for the single-polarization case, are extended to the strong coupling multi-mode case. A short recapitulation of the studied topics and a short outlook on possible future topics in chapter 6 conclude the thesis. In the appendix, a lot of additional useful information is given. In appendix A, proofs fo many relations from chapters 4 and 5 are developed in detail. Additional algorithms and tables are given in appendix B and C respectively. Many simulation results that were obtained during the studies for this thesis were cut from the main chapters for the sake of brevity. Some of these additional simulation results are given and discussed in appendix D.

2
Preliminaries

Before reviewing some elementary aspects of fiber-optic communication, definitions for some basic quantities, as well as some generalities regarding notation will be given. Additionally, useful results from the fields of information theory, operator theory and perturbation theory will be reviewed. Concluding this preliminary chapter, several evaluation metrics used in later sections of this thesis are discussed briefly.

2.1. Definitions and Notation

Due to the broad field of topics covered, adhering to the same notation style throughout the whole thesis is not always possible. Nonetheless, the thesis is conforming to the notation style declared in this section as strictly as possible and deviations are explicitly stated in the respective sections. Notation that is specific to a certain topic, is defined when the concepts are introduced over the course of the thesis.

2.1.1. Special Sets and Subsets

Frequently used sets are the integer numbers $\mathbb{Z}$, the real numbers $\mathbb{R}$ and the complex numbers $\mathbb{C}$. Each complex number has the form $z = \Re\{z\} + \mathrm{j}\Im\{z\}$, where $\mathrm{j} = \sqrt{-1}$ is the *imaginary unit* and $\Re\{\cdot\} \in \mathbb{R}$ and $\Im\{\cdot\} \in \mathbb{R}$ are the *real* and *imaginary part* respectively. Further for $z = a + \mathrm{j}b$ with $a, b \in \mathbb{R}$ the *complex conjugate* $(\cdot)^*$ is defined as $z^* = a - \mathrm{j}b$.

Common subsets of the presented sets are indicated by using sub- and superscripts. For integer numbers the set $\mathbb{Z}^+$ contains all non-negative numbers, excluding zero, and $\mathbb{Z}^-$ is the set of all negative integers, also not including zero. To include zero, the sets are written as $\mathbb{Z}_0^+$ and $\mathbb{Z}_0^-$ respectively. For the complex set $\mathbb{C}$, commonly used subsets are the positive complex half-plane $\mathbb{C}^+$, which is the set of all complex values z with imaginary part $\Im\{z\} > 0$, and the negative complex half-plane $\mathbb{C}^-$, which is the set of all complex values z with imaginary part $\Im\{z\} < 0$. Since the remaining real-line of the complex plane is already denoted by $\mathbb{R}$, no additional notation for this remaining subset is needed.

2.1.2. Vectors, Matrices and Operators

Vectors are denoted by lower-case italic bold letters, e.g. $\boldsymbol{x}$ and are assumed to be column vectors if not mentioned otherwise. Column vectors are defined as $\boldsymbol{x} = (x_1, \cdots, x_N)^{\mathrm{T}}$, where x_n with $n \in [1, N]$ are the scalar entries of the vector at position n and $(\cdot)^{\mathrm{T}}$ denotes vector transposition.

Matrices are denoted by uppercase italic bold letters e.g. $\boldsymbol{X}$. The scalar entries of matrix $\boldsymbol{X}$ are represented by $x_{i,j}$ with row-number $i \in [1, M]$ and column-number $j \in [1, N]$, where M is the number of rows and N is the number of columns of the matrix. The transposed matrix $\hat{\boldsymbol{X}} = \boldsymbol{X}^{\mathrm{T}}$ has entries $\hat{x}_{i,j} = x_{j,i}$. Further, the *Hermetian* $(\cdot)^{\mathrm{H}}$ of a matrix is defined as $\hat{\boldsymbol{X}} = \boldsymbol{X}^{\mathrm{H}} = (\boldsymbol{X}^*)^{\mathrm{T}}$ and thus its entries $\hat{x}_{i,j} = x_{j,i}^*$ are the complex conjugate entries of its transpose. Matrix dimensions are only given explicitly if instructive, using the notation $\boldsymbol{X}_{M\times N}$, to denote a matrix with M rows and N columns. This notation can be adopted for vectors, e.g. $\boldsymbol{x}_{1\times N}^{\mathrm{T}} = \boldsymbol{x}_{N\times 1}$ if necessary. This notation is also used to describe sets of matrices, e.g. $\mathbb{C}^{M\times N}$, which is the set of all matrices of dimensions $M \times N$ with complex entries.

Operators are denoted using uppercase non-italic letters e.g. X. If an operator has a matrix-structure, this is denoted by using uppercase bold non-italic letters e.g. $\mathbf{X}$.

2.1.3. Subscripts and Superscripts

Generally, non-italic sub- and superscripts are used descriptively, e.g. to distinguish the value at transmitter (TX) side x_{TX} from the value at receiver (RX) side x_{RX}. If italic letters are used for subscripts, they generally have an indexing function. In special cases, italic subscripts can also be a shorthand notation for the derivative

according to the subscript, e.g. $x_{yy} = \mathrm{d}^2x(y)/\mathrm{d}y^2$. To avoid confusion, use of this shorthand notation will be explicitly stated, if not obvious from context. Superscripts in italic are mostly used for standard power notation. One common exception is bracketed superscript notation, e.g. $x^{(i)}$, which denotes the value of parameter x at iteration i of some iterative process. If necessary subscript types are used jointly, e.g. $x_{\mathrm{TX},i}$, while clarity when handling variables with multiple superscript types will often be maintained by use of brackets, e.g. $(x^{(i)})^2$.

2.1.4. Functions and Transformations

Frequently used functions are given a (symbolic) shorthand. Commonly used functions are the absolute value $|\cdot|$, the argument of a complex number $\angle(\cdot)$. The commonly used Fourier transform $\mathcal{F}\{\cdot\}$ is defined as [34]

$$G(f) = \mathcal{F}\{g(\tau)\} = \int_{-\infty}^{\infty} g(\tau)\mathrm{e}^{-\mathrm{j}2\pi f\tau}\mathrm{d}\tau, \tag{2.1}$$

where e is Euler's number, f is the frequency parameter, τ is the time parameter and $g(\tau)$ is some time-domain signal. Its inverse $\mathcal{F}^{-1}\{\cdot\}$ is defined as [34]

$$g(\tau) = \mathcal{F}^{-1}\{G(f)\} = \int_{-\infty}^{\infty} G(f)\mathrm{e}^{\mathrm{j}2\pi f\tau}\mathrm{d}f. \tag{2.2}$$

Due to the focus on the NFT in this thesis, the transformations given in Eqs. (2.1) and (2.2) will be called the 'linear' or 'standard' Fourier transform where needed for the sake of clarity. *Parseval's theorem* relates the energy E in time domain to the energy of the Fourier transform by [34]

$$E = \int_{-\infty}^{\infty} |g(\tau)|^2\mathrm{d}\tau = \int_{-\infty}^{\infty} |G(f)|^2\mathrm{d}f. \tag{2.3}$$

Similarly, the average signal power P is defined as

$$P = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |g(\tau)|^2\mathrm{d}\tau. \tag{2.4}$$

For all the above relations it is assumed that $g(\tau)$ is absolutely and square integrable, respectively.

The *Hilbert transformation* $\mathcal{H}\{\cdot\}$ is defined as [35]

$$\mathcal{H}\{f(x)\}(y) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(x)}{y-x} \mathrm{d}x, \tag{2.5}$$

with its inverse being $\mathcal{H}^{-1}\{\cdot\} = -\mathcal{H}\{\cdot\}$. The Hilbert transform is applied in a wide variety of problems in mathematics and physics. In this thesis, its main use is to establish a relation between magnitude and phase of an analytic complex function [23].

2.1.5. Random Variables

If not stated otherwise, the following is taken from [36], [37]. A random variable (RV) X with finite sample space $\mathcal{X}$ has realizations $a \in \mathcal{X}$, which occur with a certain probability $P_X(a) = \Pr[X = a]$. The *joint probability distribution* for the two random variables X, Y is denoted by $P_{XY}(a, b)$ where $b \in \mathcal{Y}$, which is the sample space of Y. *Conditional probability* is denoted by $P_{Y|X}(b|a)$. The *mean* of a random variable is

$$\mu = \mathbb{E}[X] = \sum_{a \in \mathcal{X}} a P_X(a), \tag{2.6}$$

where $\mathbb{E}[\cdot]$ is the expected value. *Variance* σ^2 or $\mathrm{Var}[\cdot]$ is defined as

$$\sigma^2 = \mathrm{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2. \tag{2.7}$$

Similarly for two parameters the *covariance* $\mathrm{Cov}[\cdot, \cdot]$ is given by

$$\mathrm{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \tag{2.8}$$

2.2. Aspects of Information Theory

In his seminal paper "The Mathematical Theory of Communication" [38] Claude E. Shannon introduced some fundamental ideas that form the foundation for the field of *information theory*. In the following, some necessary information theoretical quantities

are presented shortly. A more in-depth treatment of information theory can be found in [39].

2.2.1. Basic Metrics

For a RV X, the amount of uncertainty about its realization is given by its *entropy*. Entropy $\mathbb{H}(\cdot)$ is computed by [38]

$$\mathbb{H}(X) = \sum_{a \in \text{supp}(P_X)} -P_X(a)\log_2\left(P_X(a)\right) = \mathbb{E}\left[-\log_2\left(P_X(a)\right)\right], \tag{2.9}$$

where $\text{supp}(\cdot)$ are all values for which $P_X(a) > 0$. Entropy is non-negative and $\mathbb{H}(X) \leq \log_2(|\mathcal{X}|)$ with $|\mathcal{X}|$ being the cardinality of $\mathcal{X}$.

The *conditional entropy* $\mathbb{H}(\cdot|\cdot)$ is obtained by

$$\mathbb{H}(Y|X) = \sum_{a \in \text{supp}(P_X)} P_X(a)\mathbb{H}(Y|X=a) = \mathbb{E}\left[-\log_2(P_{Y|X}(Y|X))\right], \tag{2.10}$$

where

$$\mathbb{H}(Y|X=a) = \sum_{b \in \text{supp}(P_{Y|X}(\cdot|a))} -P_{Y|X}(b|a)\log_2[P_{Y|X}(b|a)]. \tag{2.11}$$

The conditional entropy is bound by $0 \leq \mathbb{H}(Y|X) \leq \mathbb{H}(Y)$. If Y is fully determined by X, the conditional entropy is zero. On the other hand, if knowledge about outcome a is not reducing uncertainty of outcome b, the conditional entropy is identical to the entropy of RV Y.

From these quantities, the mutual information (MI) $\mathbb{I}(\cdot;\cdot)$ can be derived. Mutual information is defined as

$$\mathbb{I}(X;Y) = \sum_{(a,b) \in \text{supp}(P_{XY})} P_{XY}(a,b)\log_2\left(\frac{P_{XY}(a,b)}{P_X(a)P_Y(b)}\right) \tag{2.12}$$

$$= \mathbb{H}(X) - \mathbb{H}(X|Y) \tag{2.13}$$

$$= \mathbb{H}(Y) - \mathbb{H}(Y|X) \tag{2.14}$$

From Eqs. (2.13) and (2.14), it can be seen that MI is symmetric $\mathbb{I}(X;Y) = \mathbb{I}(Y;X)$. A common interpretation is that MI gives the reduction of uncertainty of one RV if one observes the other. MI is strictly nonnegative $\mathbb{I}(X;Y) \geq 0$ where $\mathbb{I}(X;Y) = 0$ if the two random variables are statistically independent.

2.2.2. Upper Bounds on Mutual Information

The maximum number of bits per symbol that can be transmitted reliably over some given channel is called the channel *capacity* and is the maximum achievable MI over all possible input distributions [40]

$$C = \max_{X} \mathbb{I}(X;Y). \tag{2.15}$$

It can be shown, that for transmitting complex symbols over the additive white Gaussian noise (AWGN)-channel, the capacity is given by [41], [39, Ch.9]

$$C_{\mathrm{AWGN}} = \log_2\left(1 + \frac{P_{\mathrm{S}}}{\sigma_{\mathrm{N}}^2}\right) = \log_2\left(1 + \mathrm{SNR}\right), \tag{2.16}$$

where P_{S} is the signal power, $P_{\mathrm{N}} = \sigma_{\mathrm{N}}^2$ is the noise power and the ratio between these powers is the signal to noise ratio (SNR).

In [13], [40], [42] it is shown that, while the exact capacity for fiber optic channels is still an open problem, an upper bound can be derived for certain scenarios. The capacity per modulated complex quantity in a multi-span point-to-point transmission case is upper-bound by [13]

$$C_{\mathrm{NLSE}} \leq \log_2\left(1 + \frac{P_{\mathrm{S,RX}}}{P_{\mathrm{N,RX}}}\right) = \log_2\left(1 + \mathrm{SNR}\right) \tag{2.17}$$

where $P_{\mathrm{S,RX}}$ is the average signal power at the receiver and the noise-power $P_{\mathrm{N,RX}}$ is the accumulated noise power at the receiver, due to e.g. amplification processes. The bound in Eq. (2.17) will be referred to as the Shannon-bound (SB) over the course of this thesis.

2.3. Aspects of Operator Theory

For the derivation of the NFT some results and definitions from the field of linear operator theory are needed and will be introduced in the following. When transitioning from finite-dimensional to infinite-dimensional spaces not all results carry over and some important new phenomena in regards to the spectrum of such an operator appear [43]. We restrict the following treatment to operators on Hilbert-spaces, since this will be the primary case over the course of this thesis. Note, that in this section $()^*$ does

denote the adjoint instead of the complex conjugate.

2.3.1. The Adjoint and Inverse Operator

Let domain and range of linear operator $\mathrm{A} : \mathcal{H}_1 \to \mathcal{H}_2$ be two Hilbert-spaces $\mathcal{H}_1$, $\mathcal{H}_2$ with inner product $\langle \cdot, \cdot \rangle$ and corresponding norm $|| \cdot || = \sqrt{\langle \cdot, \cdot \rangle}$ [44]. Operator $\mathrm{A}^* : \mathcal{H}_2 \to \mathcal{H}_1$ is the adjoint if [45]

$$\langle \mathrm{A}^* \psi, \phi \rangle = \langle \psi, \mathrm{A}\phi \rangle, \quad \forall \phi \in \mathcal{H}_1 \tag{2.18}$$

with $\psi \in \mathcal{H}_2$. If $\mathrm{A} = \mathrm{A}^*$, the operator is called *self-adjoint*.

Each bounded self-adjoint operator A is unitarily equivalent to $\Lambda = \mathrm{U}^{-1}\mathrm{AU}$, where U is some unitary operator and Λ is a *multiplication operator*. A bounded operator is an operator which always maps bounded inputs to bounded outputs. As an analogy to matrix calculus Λ often is also called the *diagonal operator*.

An operator A has an *inverse* operator A^{-1} if the domain of each operator is the range of the other and $\mathrm{AA}^{-1} = \mathrm{A}^{-1}\mathrm{A} = \mathrm{I}$, where I is the identity operator. Assuming the domain and range of operator A is the whole Hilbert space $\mathrm{A} : \mathcal{H} \to \mathcal{H}$ it is invertible if it is bijective (one-to-one and onto), and has a bounded inverse [46].

2.3.2. The Spectrum of an Operator

The spectrum $\sigma(\cdot)$ of an operator is given by [43], [44]

$$\sigma(\mathrm{A}) = \{\lambda \in \mathbb{C} | \mathrm{A} - \lambda \mathrm{I} \text{ is not invertible}\}. \tag{2.19}$$

Over the course of this thesis this spectrum will be partitioned into two different regions, depending on the cause of $(\mathrm{A} - \lambda I)$ being not invertible. If the range of operator $(\mathrm{A} - \lambda I)$ is a subset of $\mathcal{H}$ it is not onto and thus it is not invertible. The set of λ for which this is the case will be called the continuous spectrum, which does not exist for finite dimensional operators [43]. The second part of the spectrum, for which $\mathrm{A} - \lambda \mathrm{I}$ is not one-to-one and thus not invertible, is called the discrete spectrum, the respective set of λ values is called *discrete eigenvalues* and the corresponding non-zero solutions v of $\mathrm{A}v = \lambda v$ are called *eigenvectors*.

2.4. Aspects of Perturbation Theory

The channel models for which the NFT is derived, are idealized versions of some more commonly considered channel models. Especially the influence of noise and fiber-loss on the NFT is currently not fully understood. *Perturbation theory* is a mathematical framework, which can be used to study the influence of small perturbations on the spectrum of linear operators [46].

Expanding some operator A as

$$\mathrm{A} = \sum_{i=0}^{\infty} \epsilon^i \tilde{\mathrm{A}}_i, \tag{2.20}$$

where $\tilde{\mathrm{A}}_0$ is the unperturbed operator and $\epsilon^i \tilde{\mathrm{A}}_i$ are the perturbations of i-th order. Assuming the perturbations are sufficiently small, the right-hand side eigenvectors and eigenvalues of such an operator can be expanded as well to be

$$\lambda = \sum_{i=0}^{\infty} \epsilon^i \tilde{\lambda}^{(i)}, \tag{2.21}$$

$$v = \sum_{i=0}^{\infty} \epsilon^i \tilde{v}^{(i)}. \tag{2.22}$$

Once again, $\tilde{\lambda}^{(0)}$ and $\tilde{v}^{(0)}$ are the unperturbed solutions and the other terms are the perturbations of i-th order respectively.

Inserting Eqs. (2.20), (2.21) and (2.22)in $\mathrm{A}v = \lambda v$ and separating the result by powers of ϵ yields

$$\epsilon^0 : \tilde{\mathrm{A}}_0 \tilde{v}_0 = \tilde{\lambda}_0 \tilde{v}_0, \tag{2.23}$$

$$\epsilon^1 : (\tilde{\mathrm{A}}_0 - \tilde{\lambda}_0)\tilde{v}_1 + (\tilde{\mathrm{A}}_1 - \tilde{\lambda}_1)\tilde{v}_0 = 0, \tag{2.24}$$

$$\epsilon^2 : (\tilde{\mathrm{A}}_0 - \tilde{\lambda}_0)\tilde{v}_2 + (\tilde{\mathrm{A}}_1 - \tilde{\lambda}_1)\tilde{v}_1 + (\tilde{\mathrm{A}}_2 - \tilde{\lambda}_2)\tilde{v}_0 = 0. \tag{2.25}$$

$$\vdots$$

It can be seen, that Eq. (2.23) recovers the unperturbed eigenvalue equation. Equations obtained for higher powers of ϵ can be used to obtain expressions for the perturbed quantities. A wide variety of expressions can be derived from the above framework.

As an example, Eq. (2.24) can be modified by applying the inner product with some

vector u_0 to both sides, obtaining

$$\left\langle u_0, (\tilde{\mathrm{A}}_0 - \tilde{\lambda}_0)\tilde{v}_1 \right\rangle = -\left\langle u_0, (\tilde{\mathrm{A}}_1 - \tilde{\lambda}_1)\tilde{v}_0 \right\rangle. \tag{2.26}$$

Using Eq. (2.18) on the left side of Eq. 2.26, the expression $\left\langle (\tilde{\mathrm{A}}_0^* - \tilde{\lambda}_0^*)u_0, \tilde{v}_1 \right\rangle$ is obtained. Now it can be seen, that if u_0 is the eigenvector of $\tilde{\mathrm{A}}_0^*$ for $\tilde{\lambda}_0^*$, the expression vanishes and thus the left side of Eq. 2.26 becomes zero. If u_0 is chosen accordingly, the definition from [45, Def. 5.1] can be used to obtain

$$\tilde{\lambda}_1 = \frac{\left\langle u_0, \tilde{\mathrm{A}}_1 \tilde{v}_0 \right\rangle}{\langle u_0, \tilde{v}_0 \rangle} \tag{2.27}$$

from the right hand side of Eq. (2.26). Similar expressions can be derived for higher order perturbations.

2.5. Performance Metrics

Depending on the studied scenario, certain metrics can be more insightful or just more easily obtainable than others. The set of metrics used for evaluation of algorithms and system setups are given in the following

2.5.1. Mean Squared Error

The mean squared error (MSE) is the expected absolute squared difference between some expected value x and its (perturbed) measurement $\tilde{x}$. It is defined as

$$\mathrm{MSE} = \mathbb{E}_{\tilde{x}}[|\tilde{x} - x|^2]. \tag{2.28}$$

The MSE always has to be interpreted in light of the scale of the expected values x. The normalized mean squared error (NMSE) as a normalized version makes comparisons between e.g. modulation alphabets with different amplitudes more straightforward. The NMSE is defined as [47]

$$\mathrm{NMSE} = \frac{\mathbb{E}_{\tilde{x}}[|\tilde{x} - x|^2]}{\mathbb{E}_x[|x|^2]}. \tag{2.29}$$

2.5.2. Bit Error-Rate

While the MSE and NMSE are useful to e.g. gauge the accuracy of NFT algorithms, these metrics are not well suited for a complete transmission system evaluation. Since it is always assumed that digitized data is transmitted, a measurement of the transmission quality in terms of the (raw) bit error-rate (BER) is more reasonable and also widely adopted in the fiber-optic community. If a system is used to transmit bit-vector $\mathbf{b}_{\mathrm{TX}}$ of length N_{b} and the recovered bits at the receiver are $\mathbf{b}_{\mathrm{RX}}$, the BER can be computed by

$$\mathrm{BER} = \frac{N_{\mathrm{b,error}}}{N_{\mathrm{b}}}, \tag{2.30}$$

where $N_{\mathrm{b,error}}$ is the measured number of bit-errors.

Coding is not considered over the course of this thesis and thus, the BERs measured are always the 'raw' BERs. In literature it is often assumed, that if a certain BER can be achieved, there is a forward error correction (FEC) code to reduce the number of bit-errors to some small prescribed value. In literature, this FEC-threshold is often assumed to be around BER$= 10^{-3}$.

There are three points to consider regarding the use of BERs measurements for NFT-aided systems and in general. The first point is, that the FEC-threshold does not account for inhomogeneously distributed error-types like e.g. burst errors. This can somewhat be mitigated by using scrambling and descrambling algorithms at the transmitter and receiver respectively. The second point is that for many NFT-based transmission systems, the influence of e.g. time-domain noise in the NFD is not fully quantified. Thus it is unclear how an optimum bit-labeling scheme for these systems would look like. In Section 4.11 the influence of different labeling schemes on an NFT-aided transmission system is studied numerically. The third point is more practical in nature. If the BER for a certain system is measured, the number of bits that is needed for stable result, depends on the BER that is to be measured. For low BERs, the number of transmitted symbols/bits can be rather high and thus impractical, depending on the simulations complexity. Note, that while the points made above should be kept in mind, the BER and minimizing bit errors is still a concern because e.g. most state-of-the-art FEC codes work on a bit-wise basis, making low BERs still a valid optimization goal.

2.5.3. Mismatched Mutual Information

While, in theory, the MI between transmitted and received symbols can be computed according to Eq. (2.12), if transition probabilities are computed numerically, the evaluation can be time-consuming. Additionally, to facilitate the computation, binning of the received values is necessary. This results in the measured value being a lower bound to the MI.

Another more easily obtainable lower-bound is mismatched mutual information (MMI) [48], which can be computed by

$$\mathrm{MMI} = \mathbb{I}(X;Y) = \mathbb{E}\left[\log_2 \frac{q_{Y|X}(b|a)}{\sum\limits_{a' \in \mathcal{X}} P_X(a') q_{Y|X}(b|a')}\right]. \tag{2.31}$$

Here X and Y are random-variables describing the channel-input and channel-output respectively. $P_X(x')$ is the probability mass function (PMF) of X and $\mathcal{X}$ is the used constellation. Function $q_{Y|X}(b|a)$ is the conditional probability determined by the chosen mismatched channel model. It can be shown, that the obtained MMI always is a lower-bound on MI [49]. The mismatched channel model was chosen to be conditionally Gaussian as described in [48] over the whole course of this thesis. As a result, using $\approx 20 \cdot |\mathcal{X}|$ symbols is often enough to stabilize the measured result.

2.5.4. Achievable Information Rate and Spectral Efficiency

Metrics MI and MMI give the information (in bits/symbol) that can be transmitted over some given channel without errors. To be able to better compare different system setups it is often useful to compute several more generalized metrics. First the achievable information rate (AIR) is defined as

$$\mathrm{AIR} = \mathrm{MMI} \cdot R_{\mathrm{Baud}} = \mathrm{MMI}/T_{\mathrm{S}}, \tag{2.32}$$

where R_{Baud} is the systems Baud-rate and T_{S} is the symbol duration. This results in a metric stating the transmittable information in bits/second.

If the bandwidth used for transmission B_{S} is factored in as well, the SE is obtained as a measure of information that can be transmitted in bits/second/Hertz. It is defined as

$$\mathrm{SE} = \frac{\mathrm{MMI}}{\mathrm{TBP}}, \tag{2.33}$$

where the time-bandwidth product (TBP) is defined as

$$\mathrm{TBP} = T_\mathrm{S} B_\mathrm{S}. \tag{2.34}$$

For a multi-user scenario, such as the WDM system described in section 3.5, the TBP can also be defined for one super-symbol consisting of N_s symbols in time as well as N_u channels in linear frequency. The SE then can be computed equivalently using

$$\mathrm{SE} = \frac{\mathrm{MMI} \cdot N_\mathrm{u} N_\mathrm{s}}{\mathrm{TBP}} = \frac{\mathrm{MMI}}{\rho_\mathrm{ME}}, \tag{2.35}$$

where ρ_ME is called the *modulation efficiency*.

Note that, in order to ensure a fair comparison between systems, the bandwidth B_S and time-duration T_S should be chosen such that the signal is confined on the temporal and linear frequency support at any time. In a WDM system for example, the bandwidth used to compute the TBP should include potential spectral broadening. It is thus often purposeful to use the bandwidth of the receiver-filter as B_S. Additionally it has to be ensured, that if e.g. a signal broadens and subsequently shrinks in time, either the maximum attained duration is considered to be T_S or the potential interference from neighboring symbols is modeled correctly. Note that in the above relations, MI and MMI can be exchanged for each other.

3

Fiber-Optic Communication Systems

Over the course of this thesis, the term *optical communication system* always implies that the TX and RX are connected by some type of optical fiber channel. For un-amplified systems, the TX and RX are just connected by a fiber, made from silica or plastic. For long-reach systems, amplified silica fiber channels are assumed. Details on the channel models used in this thesis, are given in sections 3.1 and 3.2 for single- and multi-mode channels respectively.

Optical amplification, during transmission is realized by two different scenarios. The first amplification schemes mitigates fiber-loss at discrete points during propagation and thus will be called *lumped amplification* in the following. The details on the models for this scheme used are given in section 3.3.1. The second amplification scheme used is presented in section 3.3.2. This *distributed* scheme, continuously amplifies the signal during propagation, by pumping the fiber with light at a different frequency utilizing the effects of stimulated Raman-scattering (SRS) constructively [50, Ch.6].

Depending on the purpose of the transmission system, TX/RX schemes exist, reaching from relatively 'simple' to more elaborate. A naive approach could be to transmit data by switching a light source on and off to transmit a logical `0` or `1` respectively. While this approach might be easy to understand, it does not harness the full potential of the fiber-optic channel. Using more elaborate schemes, such as WDM, enabling multiplexing in the frequency domain and *coherent detection*, enabling the additional retrieval of the signals phase, which then in turn can be modulated, help to utilize the channel more strongly [50]. A brief review of a state-of-the-art WDM system is given

in section 3.5.

Closing this chapter on the basics of optical communication, aspects of system layouts will shortly be reviewed. Over the course of this thesis, two different scenarios are frequently assumed. The first scenario assumes that the TX and RX are connected in a *point-to-point* setup. This implies, that all spectral components propagate from the same source to the same destination. Secondly, one can also assume a generic *network-scenario*, in which many different TX and RX pairs might use the same channel. Due to the routing of the different data-streams, the common assumption is made that the RX has no knowledge about data-streams that have (partially) co-propagated with the channel of interest (COI) in the system. While the point-to-point scenario is easily understood by studying the channel and the used TX/RX structures, the network scenario and its restrictions are briefly defined in section 3.5.3.

3.1. The Single-Mode Fiber Channel

The first type of fiber channel that will be presented is the *single-mode fiber channel* or *scalar fiber channel*. It consists of spans of SMF, which are constructed such that only one fundamental mode is guided. Depending on the exact model, some amplification scheme might be deployed at the end of each fiber span. While the SMF guides two orthogonal polarizations, the mathematical model of this channel considers only one polarization, which is assumed to be orthogonal to the other polarization during transmission [14, Ch. 5]. Thus only one signal, defined by a scalar function is considered.

This channel model can be derived from the Maxwell equations. Since the derivation is rather lengthy and can be found in most of the well known books on nonlinear optics, the derivation is skipped for the sake of brevity. The channel model is given by the NLSE and can be written as [51]

$$\frac{\partial Q(\tau,\ell)}{\partial \ell} = -\frac{\alpha}{2}Q(\tau,\ell) + \frac{\mathrm{j}\beta_2}{2}\frac{\partial^2 Q(\tau,\ell)}{\partial \tau^2} - \mathrm{j}\gamma\left|Q(\tau,\ell)\right|^2 Q(\tau,\ell). \tag{3.1}$$

The parameter $Q(\tau,\ell)$ is the slowly varying signal envelope propagating in the fiber, τ is retarded time and $0 \leq \ell \leq \mathcal{L}$ is the distance the signal propagated, assuming the transmitter to be at $\ell = 0$ and the receiver to be at $\ell = \mathcal{L}$. The fiber parameters α, β_2 and γ are governing the effects of fiber-loss, dispersion and Kerr-nonlinearity, respectively.

The different terms on the right-hand side of Eq. (3.1) can be understood by their influence on the signal during propagation. These effects will be discussed in detail in the following sections. Note, that the model used here neglects several effects, such as higher order dispersion terms and stimulated scattering effects. For a more in-depth treatment of the channel model and a thorough derivation, we refer the interested reader to one of the following sources [14], [50], [51].

3.1.1. Fiber-Loss

The first right-hand side term in Eq. (3.1) is modeling the attenuation the signal experiences during propagation along the SMF. Neglecting the dispersion and nonlinearity term, Eq. (3.1) can be solved to obtain

$$Q(\tau, \ell) = Q(\tau, 0) \cdot \mathrm{e}^{\frac{-\alpha}{2}\ell}, \tag{3.2}$$

assuming that α is independent of ℓ. Attenuation coefficient α in general is not independent of frequency, thus different frequency components are attenuated differently. However, in this thesis the common assumption is made that, for transmission in the C-band (1530nm to 1565nm), the attenuation coefficient is constant [51, Ch.1.2.2].

While, as mentioned in section 1, modern optical fiber is a medium with comparably low loss, the remaining attenuation, stemming from material absorption and Rayleigh scattering, still poses certain limits on optical transmission systems. While for state-of-the-art systems utilizing e.g. WDM, the attenuation is mainly determining the amplifier spacing of the channel, attenuation strongly limits the applicability of NFT-aided systems, as will be discussed in the respective chapters of this thesis.

Using the definition of signal power from Eq. (2.4) and Eq. (3.2) the attenuation in terms of power can be given as

$$P_{\mathrm{out}} = P_{\mathrm{in}} \cdot \mathrm{e}^{-\alpha\mathcal{L}}, \tag{3.3}$$

where P_{in} and P_{out} are the power of the signal at the beginning and end of a channel of length $\mathcal{L}$ respectively.

While α has linear units [1/km] in the above equations, the attenuation is often given in [dB/km]. Conversion between the two representations can be done using [51,

Ch.1.2.2]

$$\alpha_{\mathrm{dB/km}} = -\frac{10}{\mathcal{L}} \log_{10}\left(\frac{P_{\mathrm{out}}}{P_{\mathrm{in}}}\right) = \frac{10}{\log(10)}\alpha_{1/\mathrm{km}}. \tag{3.4}$$

3.1.2. Dispersion

The second right-hand side term in Eq. (3.1) is modeling group-velocity dispersion (GVD). More generally, dispersion is caused by the frequency dependence of the effective refractive index $n(\omega)$ of the silica fiber. Thus, even if loss and nonlinearities are neglected, the signal is modified during propagation according to [52]

$$\hat{Q}(\omega, \ell) = \hat{Q}(\omega, 0) \cdot \mathrm{e}^{-\mathrm{j}\beta(\omega)\ell}, \tag{3.5}$$

where the signal is now given as its Fourier transform $\hat{Q}(\omega, \ell) = \mathcal{F}\{Q(\tau, \ell)\}$ and ω is the angular frequency. It can be seen, that each frequency component experiences a phase-shift due to propagation in ℓ according to frequency dependent mode-propagation constant $\beta(\omega)$ [51, Ch. 1.2.3].

Parameter $\beta(\omega)$ can be expanded in a Taylor-series around the frequency of the optical carrier ω_{c}, assuming that the propagating signals bandwidth is much smaller than the carrier frequency. This expansion has the form [53]

$$\beta(\omega) = n(\omega)\frac{\omega}{c} = \sum_{k=0}^{\infty} \frac{1}{k!}\beta_k(\omega - \omega_{\mathrm{c}})^k, \tag{3.6}$$

where c is the speed of light in vacuum. The coefficients β_k are defined as [53]

$$\beta_k = \left.\frac{\mathrm{d}^k\beta(\omega)}{\mathrm{d}\omega^k}\right|_{\omega=\omega_{\mathrm{c}}}. \tag{3.7}$$

The parameter β_0 represents the carriers phase-delay, which is commonly omitted since, apart from questions of e.g. TX/RX synchronization it has no significance.

Parameters β_1 and β_2 can be given explicitly as

$$\beta_1 = \frac{1}{v_{\mathrm{g}}} = \frac{n_{\mathrm{g}}}{c} = \frac{1}{c}\left(n(\omega_{\mathrm{c}}) + \omega_{\mathrm{c}} \left.\frac{\mathrm{d}n(\omega)}{\mathrm{d}\omega}\right|_{\omega=\omega_{\mathrm{c}}}\right), \tag{3.8}$$

$$\beta_2 = \frac{1}{c}\left(2 \left.\frac{\mathrm{d}n(\omega)}{\mathrm{d}\omega}\right|_{\omega=\omega_{\mathrm{c}}} + \omega_{\mathrm{c}} \left.\frac{\mathrm{d}^2n(\omega)}{\mathrm{d}\omega}\right|_{\omega=\omega_{\mathrm{c}}}\right), \tag{3.9}$$

by inserting the relation between mode-propagation constant and refractive index from

Eq. (3.6) into (3.7). It can be seen that β_1 is directly linked to the *group-velocity* v_g and *group refractive index* n_g. Thus, it can be removed from the propagation equation by defining a reference frame for the time moving at group velocity ($\tau = \tau' - \beta_1 \ell$). This is also the reason why β_1 is not present in Eq. (3.1). Parameter β_2 is the GVD parameter and is often given in form of the dispersion parameter defined as

$$D = \frac{\mathrm{d}\beta_1}{\mathrm{d}\lambda} = \frac{2\pi c}{\lambda^2}\beta_2 = \frac{\lambda}{c}\frac{\mathrm{d}^2 n(\omega)}{\mathrm{d}\lambda}, \tag{3.10}$$

where λ is the wavelength. It is often assumed, that the dispersion parameter is frequency independent unless transmission pulses are very short [14]. Similarly, higher order coefficients ($k \geq 3$) in Eq. (3.6) can be neglected as long as the pulses transmitted are not too short and transmission in not done in close proximity to the *zero-dispersion* wavelength, for which $\beta_2 \approx 0$ and $D \approx 0$. This assumptions will be made over the course of the thesis, which leads to the linear part of the channel model given by Eq. (3.1).

In general, GVD will cause phase-distortions (see Eq.(3.5)), that result in temporal pulse broadening. We would like to note that, while the term chromatic dispersion (CD) is often used to describe the same effect, we use CD as a name for a broader group of effects, including all dispersive effects due to frequency dependent propagation speeds. This is done to distinguish these effects from dispersive effects caused by other physical mechanisms e.g. polarization mode dispersion as discussed in section 3.1.4.

3.1.3. Kerr-Nonlinearity

The origin of the Kerr-nonlinearity is the real part of the *cubic material polarization* $\chi^{(3)}$. The Kerr-nonlinear effect can be divided into subgroups of perturbations, depending on the involved components in the frequency spectrum. In time domain, these effects are represented by a single term, which includes the instantaneous signal power $|Q(\tau, \ell)|^2$ and the *nonlinear parameter* γ (compare Eq. (3.1)). In this thesis, nonlinearity is treated as a monolithic effect, that potentially leads to perturbations. Thus, details on the distinguishable phenomena in the frequency domain, namely self-phase modulation (SPM), cross-phase modulation (XPM) and four wave mixing (FWM) are skipped. The interested reader is referred to [14], [50]–[52] for further details. Over the course of this thesis, it is assumed that the nonlinear response of the fiber material is instantaneous, which in general is not true, but is a reasonable assumption, as long

as the transmission pulses are not too short [51].

Similar to the treatment in section 3.1.2, Eq. (3.1) can be studied, isolating the nonlinear effects. The resulting channel model is given by

$$\frac{\partial Q(\tau,\ell)}{\partial \ell} = -\mathrm{j}\gamma \left|Q(\tau,\ell)\right|^2 Q(\tau,\ell), \tag{3.11}$$

This model has the analytic solution

$$Q(\tau,\ell) = Q(\tau,0)\cdot \mathrm{e}^{-\mathrm{j}\gamma\int_0^{\ell}|Q(\tau,\ell')|^2\mathrm{d}\ell'}. \tag{3.12}$$

Assuming that fiber loss is reincorporated into Eq. (3.11) and no (re-)amplification is present, Eq. (3.12) can be simplified further, using Eq. (3.3). The obtained solution is

$$Q(\tau,\mathcal{L}) = Q(\tau,0)\cdot \mathrm{e}^{-\mathrm{j}\gamma|Q(\tau,0)|^2\mathcal{L}_{\mathrm{eff}}}, \tag{3.13}$$

where $\mathcal{L}_{\mathrm{eff}} = (1-\mathrm{e}^{-\alpha\mathcal{L}})/\alpha$. The *effective length* $\mathcal{L}_{\mathrm{eff}}$ is the length of a hypothetical lossless fiber inducing the same amount of nonlinear distortions as its non-zero loss counterpart [51]. It is a measure for the length over which most of the nonlinear impairments in a lossy fiber are generated and will be used in section 4.9, to mitigate the degradation of NFT-generated signals in lossy channels. Curves for the relation between span-length, attenuation coefficient and effective length are given in Fig. 3.1. It can be seen, that, due to fiber loss, the main portion of nonlinear perturbations occur over the first few kilometers of the fiber span, while the contribution of the later section of the fiber is negligible, especially for higher values of α.

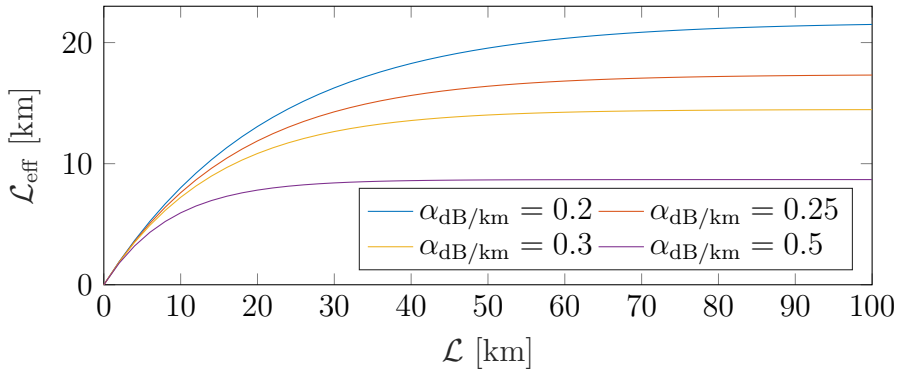


Figure 3.1.: Effective length of a passive fiber span for varying α

Also note, that the results above still hold for certain channel models which incorporate amplification. This will become more obvious in the review of amplification schemes in section 3.3.

3.1.4. Polarization Dependent Effects

As stated in the beginning of section 3.1, single-mode fibers actually guide two polarization modes. While this was neglected in the discussion of Eq. (3.1), it has to be considered in the study of PMD, which originates from the two polarizations propagating under different conditions.

If it is assumed, that the electric fields E_x, E_y of the respective polarizations propagate linearly through an idealized standard single mode fiber (SSMF), this can be written as [54]

$$\boldsymbol{E}(\ell,\tau) = \begin{pmatrix} E_x(\ell=0) \\ E_y(\ell=0) \end{pmatrix} \mathrm{e}^{-\frac{\alpha(\omega)}{2}\ell}\mathrm{e}^{-\mathrm{j}\beta(\omega)\ell}\mathrm{e}^{\mathrm{j}\omega\tau}. \tag{3.14}$$

For an ideal fiber the propagation coefficients β_x, β_y for the x-, y-polarization mode are identical ($\beta_x(\omega) = \beta_y(\omega) = \beta(\omega)$). In this case the two modes are called *degenerate*.

The polarization state of a lightwave usually is evaluated for a fixed distance ℓ and is given by the curve of the vector

$$\begin{aligned} &\Re\{\boldsymbol{E}(\ell,\tau)\} \\ =&\Re\left\{\begin{pmatrix} |E_x(\ell=0)|\mathrm{e}^{\mathrm{j}\delta_x(\ell=0)} \\ |E_y(\ell=0)|\mathrm{e}^{\mathrm{j}\delta_y(\ell=0)} \end{pmatrix} \mathrm{e}^{\mathrm{j}(\omega\tau-\beta(\omega)\ell)}\mathrm{e}^{-\frac{\alpha(\omega)}{2}\ell}\right\} \\ =&\begin{pmatrix} |E_x(\ell=0)|\cos(\omega\tau-\beta(\omega)\ell+\delta_x(\ell=0)) \\ |E_y(\ell=0)|\cos(\omega\tau-\beta(\omega)\ell+\delta_y(\ell=0)) \end{pmatrix} \mathrm{e}^{-\frac{\alpha(\omega)}{2}\ell} \end{aligned} \tag{3.15}$$

draws for varying τ [54]. Here $\delta_x(\ell=0)$, $\delta_y(\ell=0)$ are the initial phases of the complex electric fields.

Real fibers are always subject to small external perturbations such as slight core-ellipticity, torsion, bending, external pressure or external electrical/magnetic fields. As a result, the aforementioned x- and y-polarizations in a SSMF now experience *birefringence* and *mode-coupling*, leading to PMD [54]. In the following, the interactions

between the two polarizations are assumed to be fully described by [54]

$$\frac{\mathrm{d}}{\mathrm{d}\ell}\begin{pmatrix} E_x(\ell) \\ E_y(\ell) \end{pmatrix} = \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix}\begin{pmatrix} E_x(\ell) \\ E_y(\ell) \end{pmatrix}, \tag{3.16}$$

where perturbation parameters N_{11}, N_{22} are representing birefringence and N_{12}, N_{21} represent mode-coupling. If those parameters are assumed to be constant, e.g. for a short piece of fiber, Eq. (3.16) can be solved analytically. From this, a solution of the form

$$\begin{pmatrix} E_\psi(\ell) \\ E_\chi(\ell) \end{pmatrix} = \begin{pmatrix} E_\psi(\ell=0)\cdot \mathrm{e}^{-\mathrm{j}\beta_\psi \ell} \\ E_\chi(\ell=0)\cdot \mathrm{e}^{-\mathrm{j}\beta_\chi \ell} \end{pmatrix}, \tag{3.17}$$

can be stated using the eigenpolarization-coordinate system, where E_ψ, E_χ are the electric fields and β_ψ, β_χ are the propagation coefficients in eigenpolarization ψ and χ respectively. Further information on how to change between the two coordinate systems is given in [54], [55].

To quantify the effect of PMD the differential group-delay (DGD) between the eigenpolarizations is used. The DGD is given by

$$\mathrm{d}\bar{\tau}_\mathrm{g}(\omega_\mathrm{c}) = \left.\frac{\mathrm{d}(\beta_\psi(\omega)-\beta_\chi(\omega))}{\mathrm{d}\omega}\right|_{\omega_\mathrm{c}} = \bar{\tau}_{\mathrm{g},\psi}(\omega_\mathrm{c}) - \bar{\tau}_{\mathrm{g},\chi}(\omega_\mathrm{c}), \tag{3.18}$$

where $\bar{\tau}_{\mathrm{g},\psi}$, $\bar{\tau}_{\mathrm{g},\chi}$ are the average group delays in the eigenpolarizations ψ and χ respectively. If it is assumed that the perturbation parameters remain constant over the distance $\mathcal{L}_\mathrm{short}$, the DGD is given by

$$\mathrm{DGD}|_{\text{short fiber}} = \mathrm{d}\bar{\tau}_\mathrm{g}(\omega_\mathrm{c})\mathcal{L}_\mathrm{short}. \tag{3.19}$$

If the transmission distance $\mathcal{L}_\mathrm{long}$ is so long, that the perturbation parameters can no longer be assumed to be constant. In this case, the DGD between the principal states of the optical fiber link is given by [51, Ch. 1.2.4]

$$\mathrm{DGD}|_{\text{long fiber}} = D_\mathrm{PMD}\sqrt{\mathcal{L}_\mathrm{long}}, \tag{3.20}$$

where D_PMD [ps/$\sqrt{\text{km}}$] is the fiber-type dependent PMD parameter. Note, that PMD can be mitigated at the receiver by e.g. digital equalization schemes [56]. For most evaluations, done for the NLSE channel in this thesis, PMD is neglected and thus Eq. (3.1) is used as a channel model primarily.

3.1.5. Numerical Solution by Split-Step Analysis

Except for limit-cases, such as the ones described in sections 3.1.2 and 3.1.3, it is, in general, hard to solve Eq. (3.1) analytically. Thus, in most cases, the propagation equation is solved numerically. Several schemes for solving Eq. (3.1) exist [51], [57], [58], with one of the most widely used being the split-step Fourier method (SSFM). This method is used for most of the simulations of fiber-optic channels in this thesis.

Asymmetric Method

The *asymmetric method* is derived by first collecting the linear and non-linear terms in Eq. (3.1) respectively, by writing

$$\frac{\partial Q(\tau,\ell)}{\partial \ell} = (\hat{D} + \hat{N})Q(\tau,\ell), \tag{3.21}$$

$$\hat{D} = -\frac{\alpha}{2} + \frac{\mathrm{j}\beta_2}{2}\frac{\partial^2}{\partial \tau^2}, \tag{3.22}$$

$$\hat{N} = -\mathrm{j}\gamma\left|Q(\tau,\ell)\right|^2, \tag{3.23}$$

where $\hat{D}$ represents the linear and $\hat{N}$ represents the nonlinear part. Linear and nonlinear effects act jointly on the signal during propagation and thus, in general can not be viewed separately. By assuming, that for a very short distance $\Delta\ell$, linear and nonlinear effects can be viewed as acting independently, their influence on the signal can be evaluated in sequence. Thus, for this short distance $\Delta\ell$, the signal propagated according to Eq. (3.1) can be approximated by

$$Q(\tau,\ell+\Delta\ell) \approx \mathrm{e}^{\Delta\ell\hat{D}} \cdot \mathrm{e}^{\Delta\ell\hat{N}} \cdot Q(\tau,\ell). \tag{3.24}$$

The linear part can be evaluated in frequency domain, using $\mathcal{F}\{\hat{D}\} = -(\alpha/2) - (\mathrm{j}\omega^2\beta_2/2)$, while the nonlinear step can be evaluated in time-domain.

Thus, for each step of length $\Delta\ell$ one forward and one inverse Fourier transform has to be computed. Commonly this is done numerically, by using the fast Fourier transform (FFT)[59] algorithm. If it is assumed, that the cost of other operations is much lower than the cost of the two FFTs, the overall complexity is governed by the complexity of the FFT. Assuming an implementation according to Cooley-Tukey [59], the complexity in terms of "operations", meaning one complex multiplication followed by one complex addition, is given by $\mathcal{O}(N\log(N))$ [59]. For the SSFM, this leads to an

overall complexity of $\mathcal{O}(2KN\log(N))$, with K being the number of steps and N being the number of samples of the transmitted signal. The asymmetrical segmentation used in this algorithm is also depicted in Fig. 3.2, assuming that the algorithm starts with the computations for the linear step. It can be shown, utilizing the Baker-Hausdorff-

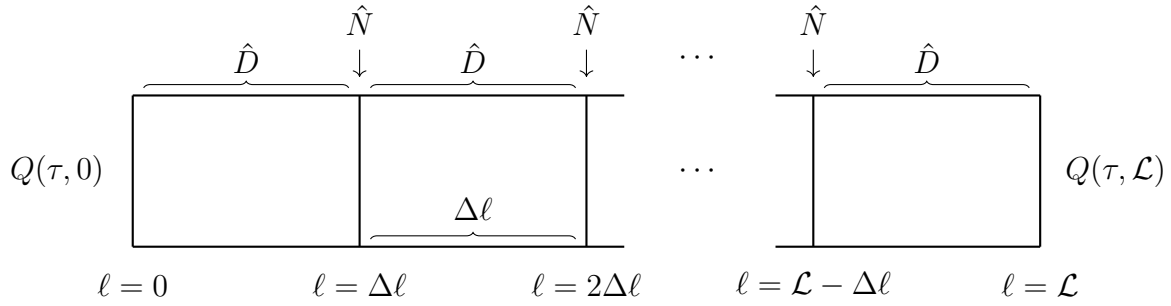


Figure 3.2.: Segmentation for the asymmetric SSFM

formula [60], that the error in each step is of second order in stepsize $\Delta\ell$ [51, ch. 2.4.1]. Since the number of steps is inversely proportional to the average stepsize, the global error of the method is $\mathcal{O}(\Delta\ell)$ [57].

Symmetric Method

The *symmetric SSFM* improves on the previously shown method, by splitting the linear computation for each step and executing the nonlinear part between the two linear sections. The signal after one step of length $\Delta\ell$ is now approximated by

$$Q(\tau, \ell + \Delta\ell) \approx \mathrm{e}^{\Delta\ell\frac{\hat{D}}{2}} \cdot \mathrm{e}^{\Delta\ell\hat{N}} \cdot \mathrm{e}^{\Delta\ell\frac{\hat{D}}{2}} \cdot Q(\tau, \ell). \tag{3.25}$$

The step-wise symmetric segmentation of the transmission distance is depicted in Fig. 3.3. It can be seen that, if $\hat{D}$ is constant in ℓ, the second linear half-step of an iteration can be executed jointly with the first linear half-step of the subsequent segment. Thus the symmetric SSFM has negligible overhead compared to the asymmetric case. The accuracy improves, such that the error is $\mathcal{O}(\Delta\ell^3)$ for the local step and $\mathcal{O}(\Delta\ell^2)$ globally [57].

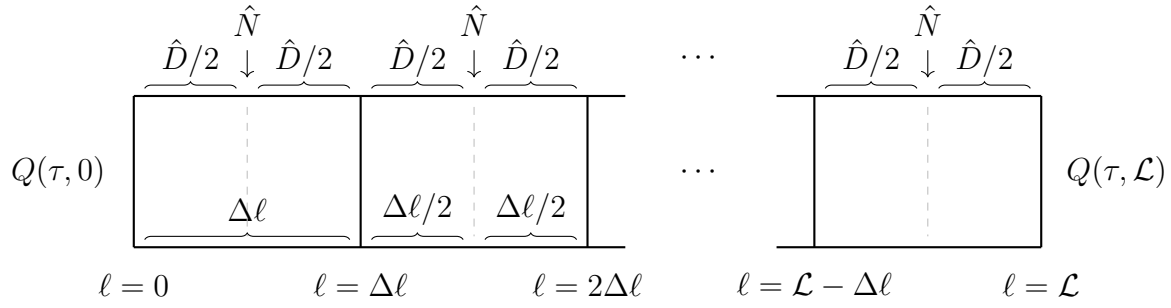


Figure 3.3.: Segmentation for the symmetric SSFM [51, ch. 2.4.1]

Step-Size Metrics

So far it was assumed, that for the numerical evaluation of Eq. (3.1) the step-size $\Delta\ell$ is known and independent of the propagated distance ℓ. By inspecting Eq. (3.1), it can be seen that varying the step-size per iteration can decrease complexity, while still achieving some target accuracy. For example, even if the initial value for $\Delta\ell$ is chosen well, the fiber-loss reduces the generated nonlinear perturbations in subsequent steps. Assuming nonlinearity is the dominant effect, the step-size could be increased to speed up the simulations without reducing the accuracy per step. In the following, three methods to set the step-size, are reviewed shortly.

Walk-Off Method The *walk-off method* is a suitable way to determine step-sizes for the SSFM if dispersion is the dominant effect during propagation [57]. This might be the case for e.g. multi-channel systems which occupy a larger bandwidth in frequency domain. The stepsize is calculated using

$$\Delta\ell = \frac{C}{|D_{\max}\lambda_{\max} - D_{\min}\lambda_{\min}|}, \tag{3.26}$$

where C is some constant to set the accuracy. $D_{\min/\max}$ and $\lambda_{\min/\max}$ are the minimum/maximum dispersion coefficients and wavelengths respectively. The denominator of Eq. (3.26) gives the difference between the slowest and fastest occurring group velocity.

Nonlinear Phase-Rotation Method If nonlinearity is the most prominent effect during propagation, the step-size can be determined using the *nonlinear phase-rotation*

method [57]. This method aims to limit the nonlinear phase-shift, that occurs due to the nonlinear part $\hat{N}$. The step-size is computed by

$$\Delta\ell \leq \frac{\phi_{\mathrm{NL}}^{max}}{\gamma \max\limits_{\tau} \left(|Q(\tau,\ell)|^2\right)}, \tag{3.27}$$

with ϕ_{NL}^{max} being the upper limit for the induced phase-shift in each step. It can be seen, that in comparison to the walk-off method, now the signal power is a determining factor for the step-size. This causes the step-size to increase for steps towards the far end of the fiber, if e.g. a passive fiber span with uncompensated fiber-loss is modeled. This can reduce the number of steps in the algorithm and speed up the numerical evaluation quite drastically.

Local Error Method The *local error method* is another adaptive method, that was first presented in [57]. For each step a 'coarse' solution $Q_{\mathrm{c}}(\tau,\ell+2\Delta\ell)$ and a 'fine' solution $Q_{\mathrm{f}}(\tau,\ell+2\Delta\ell)$ are computed. The coarse solution is computed by a single step of size $2\Delta\ell$, while the fine solution is computed by two steps of length $\Delta\ell$.

Since the exact solution is unknown, the difference between coarse and fine solution is used as a substitute measure for the local error, assuming that $Q_{\mathrm{f}}(\tau,\ell+2\Delta\ell)$ is close to the real solution. This measure is called *relative local error* and is given by

$$\delta_{\mathrm{rle}} = \frac{||Q_{\mathrm{f}}(\tau,\ell+2\Delta\ell)|| - ||Q_{\mathrm{c}}(\tau,\ell+2\Delta\ell)||}{||Q_{\mathrm{f}}(\tau,\ell+2\Delta\ell)||}. \tag{3.28}$$

Upper bounding the relative local error by δ_{B}, the algorithm distinguishes several cases and potentially updates the step-size:

- $\delta_{\mathrm{rle}} > 2\delta_{\mathrm{B}}$: Discard result, $\Delta\ell \to \frac{\Delta\ell}{2}$
- $\delta_{\mathrm{B}} \leq \delta_{\mathrm{rle}} \leq 2\delta_{\mathrm{B}}$: Step valid, $\Delta\ell \to \frac{\Delta\ell}{2^{1/3}}$
- $\frac{\delta_{\mathrm{B}}}{2} \leq \delta_{\mathrm{rle}} \leq \delta_{\mathrm{B}}$: Step valid, $\Delta\ell \to \Delta\ell$
- $\delta_{\mathrm{rle}} < \frac{\delta_{\mathrm{B}}}{2}$: Step valid, $\Delta\ell \to \Delta\ell \cdot 2^{1/3}$

If an executed step is valid, the result $Q(\tau,\ell+2\Delta\ell)$ is computed by

$$Q(\tau,\ell+2\Delta\ell) = \frac{4}{3}Q_{\mathrm{f}}(\tau,\ell+2\Delta\ell) - \frac{1}{3}Q_{\mathrm{c}}(\tau,\ell+2\Delta\ell). \tag{3.29}$$

Details on why computing this result is more optimal than just using the fine solution Q_f are given in [57].

Contrary to other methods, no initial assumptions about the strength of dispersive and nonlinear effects have to be made. Also this method does not include a way to meaningfully set the initial value for step-size $\Delta\ell$. The initial choice of $\Delta\ell$ does not significantly influence the accuracy of the local error method, since steps with a bad relative local accuracy are repeated, but a poorly chosen initial value can cause some computational overhead for the first few steps. Due to the computation of two solutions in every step, there also is a general computational overhead compared to the other methods presented in this section. In return the method exhibits a reduced global error of $\mathcal{O}(\Delta\ell^3)$.

3.2. The Strong Coupling Multi-Mode Fiber Channel

As noted in section 3.1, the scalar NLSE is a channel model often utilized to describe the most important effects in a SSMF. As implied by its name, the SSMF is only guiding a single mode. By changing the fiber geometry, fibers which guide more than one mode can be realized. Even though, the propagation in such a multi-mode fiber (MMF) is generally involved, some assumptions and restrictions can be applied to arrive at a case, for which the NFT framework can be utilized. The assumption, made over the course of this thesis is that the excited modes are strongly coupling with each other during propagation and thus the channel model is the strong-coupling Manakov equation (SCME) and transmission is taking place in the *strong-coupling regime* [61]. The resulting multi-mode channel model is given by [61], [62]

$$\frac{\partial \boldsymbol{Q}(\tau,\ell)}{\partial \ell} = -\frac{\alpha}{2}\boldsymbol{Q}(\tau,\ell) + \mathrm{j}\frac{\bar{\beta}_2}{2}\frac{\partial^2 \boldsymbol{Q}(\tau,\ell)}{\partial \tau^2} - \mathrm{j}\gamma\kappa\sum_{m=1}^{M}|Q_m(\tau,\ell)|^2\boldsymbol{Q}(\tau,\ell). \tag{3.30}$$

The vector $\boldsymbol{Q}(\tau,\ell)$ holds the signals $Q_m(\tau,\ell)$, propagating in modes $m \in 1,\cdots M$, where M is the number of stimulated modes. It can be seen, that, similar to the dual-polarization NLSE [51, Ch. 6.6.3], the co-propagating modes influence each other during propagation, by nonlinear coupling. The nonlinear coupling coefficient κ gives the factor, by which γ is altered, due to this interactions. The exact value for κ depends on the specific fiber-type used. The parameter $\bar{\beta}_2$ is the average GVD across all stimulated modes, which is also assumed to be the GVD parameter for each of the co-

propagating modes. While some strong assumptions are made to arrive at the SCME channel model, these assumptions are also commonly made in other literature [63]–[65]. With some of the neglected perturbations being reintroduced in later simulations, the lossless version of Eq. (3.30) will be the central channel model, for which the NFT and NFT-aided transmission schemes are studied.

The choice of the nonlinear coupling coefficient κ is rather involved and a detailed treatment of its computation is deemed out of scope for this thesis. The interested reader is referred to [61], [66] for further details. The exact value of κ, in most cases, is not relevant for the topics studied in this thesis. Thus, the approximation $\kappa = \frac{8}{9}\frac{2}{M}$ [66] is used if not stated otherwise. Note that, while in [66] and other publications the notion of spatial modes is adopted, the term mode will refer to one polarization in a spatial mode over the course of this thesis, making e.g. the dual-polarization NLSE a two mode system.

3.3. Amplification Schemes

Compared to other transmission media, the loss in modern optical fiber is relatively low. Nonetheless, depending on the transmission distance and transmission scheme used, it still has a significant influence on the transmission quality and thus has to be mitigated. In this thesis, two main types of amplification are utilized. The first type is *lumped amplification*, meaning amplification that is conducted at discrete points in the transmission link. The prime example for such schemes is amplification by EDFAs, which will be reviewed shortly in section 3.3.1. The second type is *distributed amplification*, which means that even though amplification units are placed at discrete points in the system, amplification of the signal potentially takes place along the whole transmission distance. Details on distributed Raman amplification (DRA) [67], as the prime example for distributed amplification schemes, is discussed in section 3.3.2.

Note that the different amplification schemes will be reviewed for the scalar NLSE case only. For the multi-mode channel described in section 3.2, it will be assumed that the amplification process per mode is identical to the scalar NLSE case.

3.3.1. Lumped Amplification by Erbium Doped Fiber Amplifiers

For lumped amplification, the fiber-channel itself is considered to be *passive*, meaning that the signal power is decreasing during propagation, according to Eq. (3.3). Then,

after a certain distance $\mathcal{L}_{\text{span}}$, the attenuated signal is amplified. The signal does not necessarily always have to be re-amplified to its initial power, but over the course of this thesis it will be assumed that the signal power is fully recovered at the end of each passive fiber segment. Those passive fiber segments will from now on be called *(fiber-)spans*. Due to the previously made assumption, the EDFA after each span has to have the amplifier gain

$$G_{\text{EDFA}} = \mathrm{e}^{\alpha \mathcal{L}_{\text{span}}}. \tag{3.31}$$

Note that, in general, a concatenation of EDFA amplified fiber spans does not necessarily have to consist of spans of equal length. Nonetheless, in this thesis, it is assumed that amplification is done periodically and a transmission channel consist of N_{span} spans of the same length $\mathcal{L}_{\text{span}}$.

Each EDFA instance is inevitably generating additive amplified spontaneous emission (ASE) noise. The noise variances for one span ($\sigma^2_{\text{ASE,span}}$), as well as for the whole fiber link (σ^2_{ASE}), are given by [68], [40], [50, Ch.7]

$$\sigma^2_{\text{ASE,span}} = (G_{\text{EDFA}} - 1) h_{\text{P}} f_{\text{c}} n_{\text{sp}} B_{\text{rec}}, \tag{3.32}$$

$$\sigma^2_{\text{ASE}} = N_{\text{span}} \sigma^2_{\text{ASE,span}}, \tag{3.33}$$

where h_{P} is the Planck constant, f_{c} is the frequency of the optical carrier, n_{sp} is the *spontaneous emission factor* and B_{rec} is the receiver bandwidth. Note that for this whole section the single-polarization NLSE is the assumed channel model and thus the expressions given in Eqs. (3.33) and (3.32) holds for the single polarization case. The noise figure can be computed according to

$$F_{\text{n}} = 2 n_{\text{sp}} (1 - \frac{1}{G_{\text{EDFA}}}) + \frac{1}{G_{\text{EDFA}}} \tag{3.34}$$

In the following it will be assumed, that the generated ASE noise is white in the frequency range of interest and that the approximation for the *noise figure* $F_{\text{n}} \approx 2 n_{\text{sp}}$ [51, Ch. 6.1.3] holds. Note, that the spontaneous emission factor is $n_{\text{sp}} = 1$ for an ideal amplifier according to [50, Ch. 7.2.3] and thus still would result in $F_{\text{n}} = 3\text{dB}$. Over the course of this thesis, the more realistic value $F_{\text{n}} = 5\text{dB}$, which is commonly found in related literature, will be used. Note, that Eq. (3.33) could be easily modified, resulting in a sum of per-span variances with differing $\mathcal{L}_{\text{span}}$ and G_{EDFA} to allow for a wider variety of setups.

3.3.2. Distributed Amplification by Raman Amplifiers

For distributed amplification, the fiber channel is used as an active part of the amplifier structure. The underlying effect used in the DRA amplification scheme is SRS, amplifying the signal due to a strong pump-beam propagating with the signal. The signal- and pump-power with respect to ℓ are described by the coupled equations [50]

$$\frac{\mathrm{d}P_\mathrm{S}}{\mathrm{d}\ell} = -\alpha_\mathrm{S} P_\mathrm{S} + \frac{g_\mathrm{R}}{a_\mathrm{P}} P_\mathrm{P} P_\mathrm{S}, \tag{3.35}$$

$$\frac{\mathrm{d}P_\mathrm{P}}{\mathrm{d}\ell} = -\alpha_\mathrm{P} P_\mathrm{P} - \frac{\omega_\mathrm{P}}{\omega_\mathrm{S}} \frac{g_\mathrm{R}}{a_\mathrm{P}} P_\mathrm{S} P_\mathrm{P}, \tag{3.36}$$

where P_x, ω_x, α_x with $\mathrm{x} \in \{\mathrm{S}, \mathrm{P}\}$ are the power, frequency and fiber attenuation for the signal and pump respectively. g_R is the Raman gain and a_P is the area of the pump-beams cross-section in the fiber. Note that the fraction $g_\mathrm{R}/a_\mathrm{P}$ is often considered to be one parameter called the *Raman-gain efficiency*. For details on modeling and computation of the Raman-gain efficiency the interested reader is referred to [69].

The pump laser can be injected into the fiber, such that it either co- or counter-propagates in relation to the signal. This results in a variation of the signals power-profile during propagation depending on the exact setup. Three different configurations for DRA will be discussed in the following.

Subsequently, a fourth DRA setup with idealized characteristics will be presented. This fourth scheme neglects residual fluctuations of the signal power and thus can be seen as ideal distributed amplification with added noise.

Co-Propagating Pump

For the first configuration, the pump-beam is injected into the fiber at the same end as the signal. Thus the signal and pump co-propagate. Assuming small-signal amplification the second term in Eq. (3.36), which models pump-depletion can be neglected. The pump power at distance $\ell \in [0, \mathcal{L}_\mathrm{span}]$ then is given by [70]

$$P_\mathrm{P}(\ell) = P_\mathrm{P}(0)\mathrm{e}^{-\alpha_\mathrm{P}\ell}. \tag{3.37}$$

Inserting Eq. (3.37) into (3.35) the signal power at distance ℓ can be expressed by

$$P_\mathrm{S}(\ell) = P_\mathrm{S}(0) \cdot \mathrm{e}^{\frac{g_\mathrm{R}}{a_\mathrm{P}} P_\mathrm{P}(0) \frac{1-\mathrm{e}^{-\alpha_\mathrm{P}\ell}}{\alpha_\mathrm{P}} - \alpha_\mathrm{S}\ell}. \tag{3.38}$$

Often it is desirable to fully compensate the attenuation due to fiber loss after each span. To achieve this the initial pump power $P_{\mathrm{P}}(0)$ has to be chosen such that, according to Eq. (3.38), $P_{\mathrm{S}}(\mathcal{L}_{\mathrm{span}}) = P_{\mathrm{S}}(0)$. As a result the initial pump power for a specific span length is given by

$$P_{\mathrm{P}}(0) = \frac{\alpha_{\mathrm{S}}\mathcal{L}_{\mathrm{span}}a_{\mathrm{P}}}{\mathcal{L}_{\mathrm{eff}}g_{\mathrm{R}}}, \tag{3.39}$$

where $\mathcal{L}_{\mathrm{eff}} = \frac{1-\mathrm{e}^{-\alpha_{\mathrm{P}}\mathcal{L}_{\mathrm{span}}}}{\alpha_{\mathrm{P}}}$ is the effective fiber length.

Counter-Propagating Pump

If the pump-beam is injected into the receiver-side end of the fiber, the pump and signal are said to be *counter-propagating*. The pump-power is now increasing with increasing ℓ, and thus Eq. (3.36) has to be modified to be

$$-\frac{\mathrm{d}P_{\mathrm{P}}}{\mathrm{d}\ell} = -\alpha_{\mathrm{P}}P_{\mathrm{P}}, \tag{3.40}$$

Note that in Eq. (3.40) pump depletion is again neglected.

The pump power at distance $\ell \in [0, \mathcal{L}_{\mathrm{span}}]$ can be given as [70]

$$P_{\mathrm{P}}(\ell) = P_{\mathrm{P}}(\mathcal{L}_{\mathrm{span}})\mathrm{e}^{-\alpha_{\mathrm{P}}(\mathcal{L}_{\mathrm{span}}-\ell)}, \tag{3.41}$$

where $P_{\mathrm{P}}(\mathcal{L}_{\mathrm{span}})$ is the initial power of the pump-beam injected at the receiver end of the fiber span. The signal power with respect to ℓ then is

$$P_{\mathrm{S}}(\ell) = P_{\mathrm{S}}(0)\mathrm{e}^{\frac{g_{\mathrm{R}}}{a_{\mathrm{P}}}P_{\mathrm{P}}(\mathcal{L}_{\mathrm{span}})\frac{\mathrm{e}^{-\alpha_{\mathrm{P}}(\mathcal{L}_{\mathrm{span}}-\ell)}-\mathrm{e}^{-\alpha_{\mathrm{P}}\mathcal{L}_{\mathrm{span}}}}{\alpha_{\mathrm{P}}}-\alpha_{\mathrm{S}}\ell}. \tag{3.42}$$

Again, the pump power to perfectly compensate loss at the end of the fiber-span can be calculated, using Eq. (3.42) and $P_{\mathrm{S}}(\mathcal{L}_{\mathrm{span}}) = P_{\mathrm{S}}(0)$. The injected pump power is then given by

$$P_{\mathrm{P}}(\mathcal{L}_{\mathrm{span}}) = \frac{\alpha_{\mathrm{S}}\mathcal{L}_{\mathrm{span}}a_{\mathrm{P}}}{\mathcal{L}_{\mathrm{eff}}g_{\mathrm{R}}}. \tag{3.43}$$

Due to symmetries, the effective fiber length $\mathcal{L}_{\mathrm{eff}}$ is computed identically for the co- and counter-propagating pump case. Also note, that the choice between co- and counter-propagating pump only influences the gain distribution along the fiber span, but does not influence the overall gain after one span. Thus, as it can be seen in Eqs. (3.39)

and (3.43), the injected pump-power has to be identical in both cases.

Even Power Bidirectional Pump

For the third configuration, it is assumed that two pump-lasers are injected into the fiber, one co- and one counter-propagating [71]. Generally, these pumps can have different wavelengths and input powers. In the case described in this section, it is assumed that both pump lasers have the same wavelength and are injected into the fiber with identical input power $P_\mathrm{P}(0) = P_\mathrm{P}(\mathcal{L}_\mathrm{span})$. Propagation of the signal and pump lasers is described by Eqs. (3.35), (3.36) and (3.40).

The total pump power at distance ℓ is given by

$$P_\mathrm{P}(\ell) = P_\mathrm{P}(0)\mathrm{e}^{-\alpha_\mathrm{P}\ell} + P_\mathrm{P}(0)\mathrm{e}^{-\alpha_\mathrm{P}(\mathcal{L}_\mathrm{span}-\ell)} \tag{3.44}$$

and the signal power at distance ℓ is given by

$$P_\mathrm{S}(\ell) = P_\mathrm{S}(0)\mathrm{e}^{\frac{g_\mathrm{R}}{a_\mathrm{P}}P_\mathrm{P}(0)\frac{1-\mathrm{e}^{-\alpha_\mathrm{P}\ell}+\mathrm{e}^{-\alpha_\mathrm{P}(\mathcal{L}_\mathrm{span}-\ell)}-\mathrm{e}^{-\alpha_\mathrm{P}\mathcal{L}_\mathrm{span}}}{\alpha_\mathrm{P}}-\alpha_\mathrm{S}\ell}. \tag{3.45}$$

The pump power for each pump-beam, needed to fully compensate the fiber loss in the span, is thus given by

$$P_\mathrm{P}(0) = \frac{\alpha_\mathrm{S}\mathcal{L}_\mathrm{span}a_\mathrm{P}}{2\mathcal{L}_\mathrm{eff}g_\mathrm{R}}. \tag{3.46}$$

The variation of the signal power during propagation depends on the span-length $\mathcal{L}_\mathrm{span}$. In Fig. 3.4 the signal power over several spans of fiber is shown for different amplification types and span-lengths. It can be seen, that, especially for the case of DRA with short spans, the power fluctuations can be kept fairly flat. The DRA results in Fig. 3.4 are for co- and counterpropagating pumps.

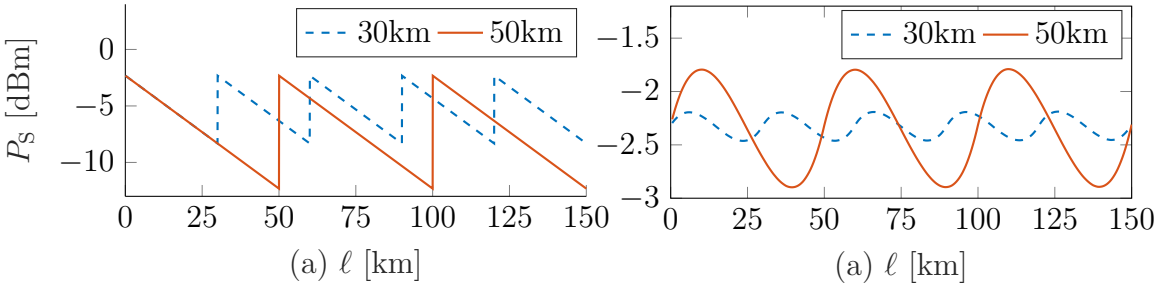


Figure 3.4.: Signal power (P_S) for (a) EDFA and (b) Raman amplification over $\mathcal{L}$ = 150km for different span lengths $\mathcal{L}_\mathrm{span}$

Idealized Raman Amplification

For idealized distributed Raman amplification (IDRA), it is assumed that fiber loss is ideally compensated at each position $\ell \in [0, \mathcal{L}]$. This effectively removes the loss term from Eq. (3.1) in exchange for distributed ASE noise. The channel model can thus be given by

$$\frac{\partial Q(\tau,\ell)}{\partial \ell} = \mathrm{j}\frac{\beta_2}{2}\frac{\partial^2 Q(\tau,\ell)}{\partial \tau^2} - \mathrm{j}\gamma Q(\tau,\ell)|Q(\tau,\ell)|^2 + N(\tau,\ell), \tag{3.47}$$

where $N(\tau,\ell)$ is the ASE noise term. In view of Fig. 3.4 it can be seen, that this idealized channel model also is the limit case of the three configurations described above for $\mathcal{L}_{\text{span}} \to 0$. For all four amplification schemes described in this section, the signal power over one fiber span is shown in Fig. 3.5 for some commonly used parameters.

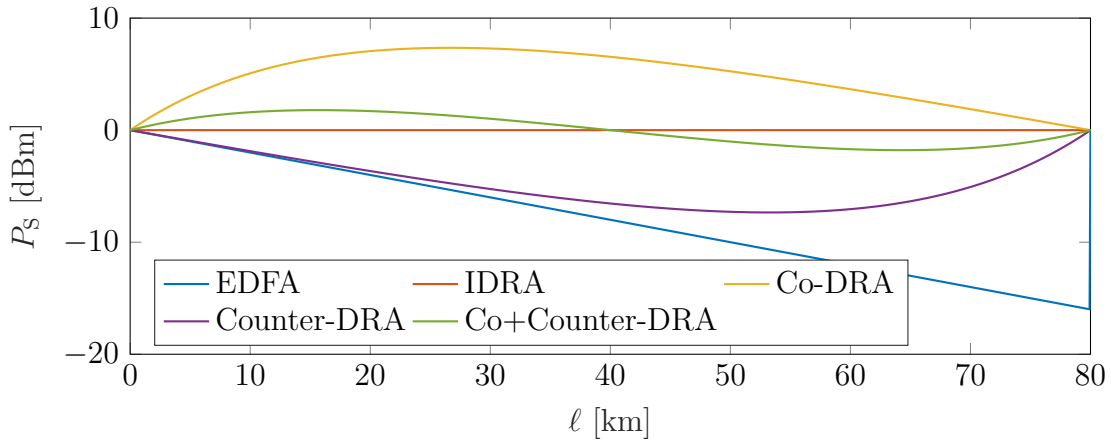


Figure 3.5.: Power profiles for EDFA and DRA amplifiers ($\mathcal{L}_{\text{span}} = 80\text{km}$, $\alpha_{\text{S}} = 0.2\text{dB/km}$, $\alpha_{\text{P}} = 0.25\text{dB/km}$, $P_{\text{S}}(0) = 0\text{dBm}$).

Noise in Raman Amplified Fibers

The IDRA channel model, is approximated by $N_{\text{span}} \to \infty$ cascaded EDFA amplified spans with vanishing span-length $\mathcal{L}_{\text{span}}$. Using Eqs. (3.31), (3.32), the relation $N_{\text{span}} = \mathcal{L}/\mathcal{L}_{\text{span}}$ and computing the limit of Eq. (3.33) for $\mathcal{L}_{\text{span}} \to 0$, the noise variance for the IDRA channel model ASE noise is found to be [40]

$$\sigma^2_{\text{IDRA}} = \alpha_{\text{S}}\mathcal{L}_{\text{span}}N_{\text{span}}h_{\text{P}}f_{\text{c}}K_{\text{T}}B_{\text{rec}} = \alpha_{\text{S}}\mathcal{L}h_{\text{P}}f_{\text{c}}K_{\text{T}}B_{\text{rec}}, \tag{3.48}$$

where $\mathcal{L} = \mathcal{L}_{\text{span}} N_{\text{span}}$ is the total transmission length. Note, that here the spontaneous emission factor was replaced using the *shifted phonon occupancy factor* $K_{\text{T}} = 1 + \eta_{\text{P}}$, where η_{P} is the *phonon occupancy factor* [71], matching the notation used in [40]. Parameter K_{T} is equivalent to the spontaneous emission factor for Raman amplified fiber n_{sp}, which now is defined as [50, Ch. 7.3.3]

$$n_{\text{sp}} = \frac{1}{1 - \exp(\frac{-h_{\text{P}}\Omega_{\text{R}}}{k_{\text{B}}T_{\text{fib}}})}. \tag{3.49}$$

Here, Ω_{R} is the *Raman shift*, which is the frequency shift between the signal and pump frequency. Further, k_{B} is the well-known Boltzmann constant. The phonon occupancy factor thus is given by

$$\eta_{\text{P}} = \frac{1}{\exp(\frac{h_{\text{P}}\Omega_{\text{R}}}{k_{\text{B}}T_{\text{fib}}}) - 1}. \tag{3.50}$$

For $\Omega_{\text{R}} \approx 13\text{THz}$, which is a value for which the Raman gain peak is reached [72], the phonon occupancy factor is $\eta_{\text{P}} \approx 0.14$ at $T_{\text{fib}} \approx 298\text{K}$ ($\approx 25°$ Celsius) [71]. Thus $n_{\text{sp}} \approx 1.14$, which matches the values given in [40].

For the three remaining schemes the computation of the noise variance can be facilitated by expressing the total noise power via the effective noise figure. If the noise power is given by [54, Ch.7],[68, Ch.2]

$$\sigma^2_{\text{DRA}} = \frac{1}{2} F_{\text{eff}} h_{\text{P}} f_{\text{c}} G_{\text{on-off}} \tag{3.51}$$

Here the parameter $G_{\text{on-off}}$ is the gain that would be necessary in an EDFA amplifier at the end of the fiber span to achieve the same amplification as the DRA amplifier used. In the case of complete reamplification at the end of the fiber span, this value is given by Eq. (3.31). This surrogate system then can be used to compute the noise power generated by the DRA according to the above equation. The *effective noise figure* F_{eff} is capturing the fact that noise, that is generated due to DRA amplification, somewhere in the fiber will be attenuated by fiber-loss and amplified by the DRA process or the remaining transmission length. It can be seen, that e.g. for backwards pumped DRA, most of the amplification takes place close to the receiver and for forward pumped DRA the main portion of amplification takes place close to the transmitter (see Fig. 3.5). Thus, in the forward pumped DRA case, the noise is attenuated for a longer propagation distance and thus is reduced in comparison to backwards pumped systems. This behavior is captured by the effective noise figure

F_{eff}. More details on the computation of F_{eff} for specific cases can be found in [71], [73]. Other ways to describe the noise accumulation in DRA amplified spans can be found in [72], [74], [50, Ch.7] and a treatment of the simulation of DRA amplifier noise is given in [75]. While the noise in DRA is often approximated as white Gaussian, this does only hold if the used bandwidth is small compared to the DRA bandwidth [72], [75]. Also if the Raman gain g_{R} is modeled as frequency dependent according to e.g. [69] this assumption does no longer hold. Over the course of this thesis, it is assumed that the signal-bandwidth is always much smaller than the DRA-bandwidth and thus the noise in DRA schemes is assumed to be white Gaussian.

3.4. Modulation Formats and Bit-Labeling

The exact physical quantity modulated in a system varies depending on the system itself. However, for most transmission schemes, a wide variety of usable *modulation formats* exists [35], [40], [76]. These can be treated independent of the physical realization of the exact system they are used in. In the following, two modulation formats, namely quadrature amplitude modulation (QAM) [35] and multi-ring phase-shift keying (MPSK) [40], will be shortly reviewed.

Even though *bit-labeling* of constellation data is not a major concern in this thesis, the impact of several labeling methods will be studied in detail in sections 4.11. Thus, some basics regarding bit-labeling are shortly reviewed in this section as well.

3.4.1. QAM Constellation

QAM is a widely used constellation in state-of-the-art communication systems. Stemming from the multiplexing of two pulse amplitude modulation (PAM) signals on orthogonal carriers, it represents data by complex constellation points a_i, for which the real part is often called the in-phase (I) component and the imaginary part is often called the quadrature (Q) component. Fig 3.6 depicts some commonly used constellations from the M-QAM family where M is the constellations cardinality.

It can be seen, that in order to obtain quadratic constellations with an integer number of labeling bits per point, the constellation size has to be $M = 2^{2i}$ for $i \in \mathbb{N}^+$. While there are many practical examples for QAM schemes that do not adhere to this restrictions [40], we restrict ourselves to these cases over the course of this thesis for the sake of simplicity. The MI for several M-QAM constellations transmitted over an

(a) $M = 4$ (b) $M = 16$

Figure 3.6.: M-QAM Constellations

AWGN channel are depicted in Fig. 3.7, together with the SB given by Eq. (2.16).

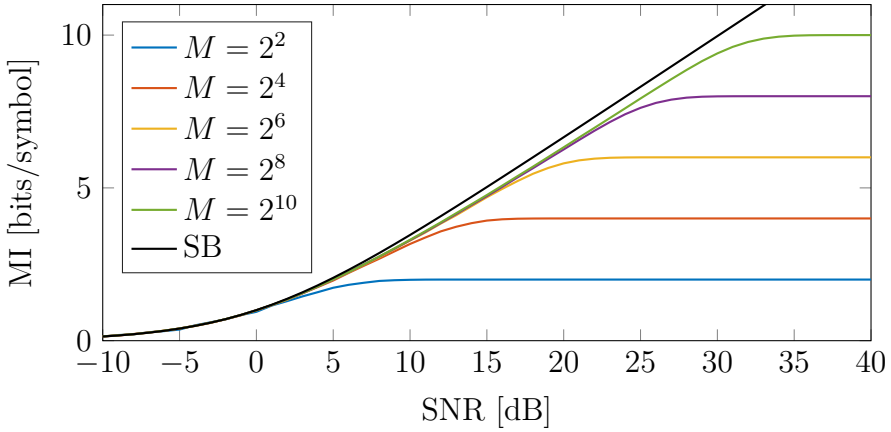
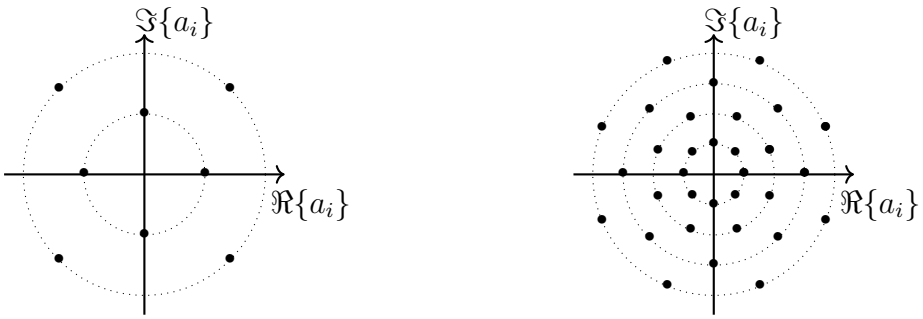


Figure 3.7.: MI for M-QAM constellation in an AWGN channel

3.4.2. MPSK Constellation

The other modulation format that will be used over the course of this thesis is MPSK, which places its constellation points a_i on rings in the complex plane. The constellations cardinality $M = N_\mathrm{r} N_\mathrm{p}$ is defined by the number of rings N_r used and the number of points per ring N_p. The MPSK constellation is constructed as an approximation of a Gaussian distribution [65]. Examples for such constellations, assuming equidistant points on equidistant rings, are given in Fig. 3.8.

The MI for several MPSK constellations transmitted over an AWGN channel, are given in Fig. 3.9.



(a) $N_\mathrm{r} = 2$, $N_\mathrm{p} = 4$ (b) $N_\mathrm{r} = 4$, $N_\mathrm{p} = 8$

Figure 3.8.: MPSK Constellations

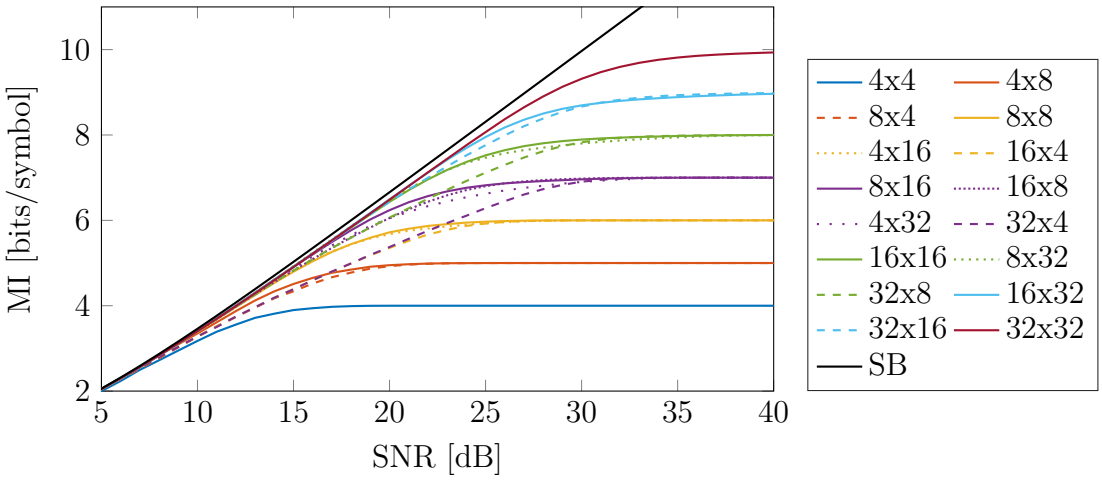


Figure 3.9.: MI for N_rxN_p MPSK constellation in an AWGN channel.

3.4.3. Bit-Labeling Schemes

Bit labeling defines the mapping of (complex) constellation points a_i onto sets of bits $\mathbf{b}_i$. Commonly either dual- or Gray-coding is used for this mapping. If it can be assumed, that the constellation points are only perturbed by circularly symmetric AWGN, then gray-coding is shown to be optimal. However, for some of the systems presented in this thesis, the noise statistics in the modulation domain are not fully known and thus some other labeling schemes might improve the BER performance. In the following, the used labeling schemes are shortly reviewed.

Gray Labeling

Gray-labeling, also called *Gray-coding*, is a widely applied scheme, used to minimize the BER in a system in relation to some given symbol error rate (SER), assuming the constellation symbols are perturbed by AWGN. Gray-labeling ensures that adjacent constellation symbols are labeled such, that their bit label only differs in one position [35, Ch.4]. Assuming that the symbols are perturbed by circularly symmetric AWGN, the most likely error occurs between adjacent symbols. Thus it makes sense to ensure that in this most likely error-cases, one symbol-error only leads to one bit-error. Note that, while not always strictly necessary, adjacent here is meant in the euclidian sense. In Fig. 3.10 two small examples for Gray-labeled constellations are shown.

(a) 4-QAM (b) 4-PAM

Figure 3.10.: Two Gray-labeled Constellations

Exhaustive Search Labeling

If no assumptions on the nature of the occurring perturbation can be made, Gray-labeling, using the euclidian distance, might no longer be optimal. *Exhaustive search labeling* is not based on any assumptions regarding the noise, other than it being stationary at least for the time the generated bit-labeling is used. However, it relies on pilot-data to generate the mapping. Exhaustive search labeling uses all possible mappings on the set of pilot-data and chooses the labeling that results in the lowest BER. Note, that this results in the optimum bit-labeling for the pilot-data, but not necessarily for the payload. The performance thus relies on the amount of pilot-data available. Additionally, the complexity of exhaustive search labeling grows rapidly with M and thus the scheme can realistically only be used for constellations with rather low cardinality. In this thesis, this method is mostly used to get some heuristic lower bounds to compare other schemes against.

Greedy Search Labeling

Another heuristic scheme is *greedy search labeling*. It again uses pilot-data to estimate the transition probabilities between constellation points. Then pairs of constellation points are labeled, such that the labels only differ in one position, starting with the pairs that have the highest transition probability. It can be seen, that this scheme might lead to some suboptimal decisions, especially in the last few iterations, since only two constellation points at a time are factored into each decision step. An example of steps that cause such a sub-optimal case is depicted in Fig. 3.11.

(a) Step 1 and 2 (b) Step 3

Figure 3.11.: Sub-optimal steps in greedy labeling (Probabilities: $P_1 > P_2 > P_3$)

There are many ways to improve the performance of this rather simple greedy-algorithm, e.g. by using branching paths to find a better mapping. However, since mapping schemes are not a main focus of this thesis, the greedy algorithm was not further augmented to improve its performance.

3.5. Wave-Division Multiplexing Transmission Systems

WDM systems are considered to be the current state-of-the-art for fiber-optic long-haul transmission [77], [78] and will be used as a baseline for comparison with transmission schemes presented in sections 4 and 5. The fundamental idea behind WDM is to multiplex signals in the frequency domain, enabling utilization of the full available bandwidth, as well as serving several users over the same fiber channel. Bandwidth itself is commonly assumed to be limited by the bandwidth of amplifiers or electro-optical converters [40], [79]. While recently there was an increased foray into ultra-wideband transmission systems, we assume that all described systems are operating in C-band.

3.5.1. Transmission and Detection

In the systems TX, constellation symbols a_{nk} are modulated onto the transmission pulse $g_{\mathrm{TX}}(\tau)$, which again is modulated such, that its spectrum is shifted into some prescribed frequency channel. This can be formulated as

$$s(\tau) = \sum_{n=0}^{N_\mathrm{s}-1} \sum_{k=1}^{N_\mathrm{u}} a_{nk} \cdot g_{\mathrm{TX}}(\tau - nT_\mathrm{S}) \cdot \mathrm{e}^{\mathrm{j}2\pi f_k \tau}, \tag{3.52}$$

where $s(\tau)$ is the transmit signal, N_s is the number of symbols transmitted per user, N_u is the number of users, a_{nk} is the n-th constellation symbol transmitted by the k-th user, $g_{\mathrm{TX}}(\tau)$ is the base-band transmission pulse, T_S is the symbol duration, and f_k is the center-frequency of the k-th users channel. While the word 'user' implies that N_u data-streams for separate TX/RX pairs co-propagate, all N_u channels can be utilized by a single TX/RX pair as well. The exact network topology does not influence the modulation/demodulation in the TX/RX. It just might impose some constraints on the mitigation steps in the RX, as will be seen in section 3.5.3.

The transmission pulse $g_{\mathrm{TX}}(\tau)$ can be chosen freely, with some being more favorable than others. In this thesis, the utilized pulse-shape for WDM is the root raised cosine (RRC). It is often described by its relation to the *raised cosine* [35, Ch. 9.2.1]. However, the RRC in time-domain can also be stated explicitly as [80]

$$g_{\mathrm{RRC}}(\tau) = \begin{cases} \frac{1}{\sqrt{T_\mathrm{S}}} \left[1 + \beta_\mathrm{RO}\left(\frac{4}{\pi} - 1\right)\right] & ,\tau = 0 \\ \frac{\beta_\mathrm{RO}}{\sqrt{2T_\mathrm{S}}} \left[\left(1 + \frac{2}{\pi}\right)\sin\left(\frac{\pi}{4\beta_\mathrm{RO}}\right) + \left(1 - \frac{2}{\pi}\right)\cos\left(\frac{\pi}{4\beta_\mathrm{RO}}\right)\right] & ,\tau = \pm\frac{T_\mathrm{S}}{4\beta_\mathrm{RO}} \\ \frac{1}{\sqrt{T_\mathrm{S}}} \frac{\sin\left[\pi\frac{\tau}{T_\mathrm{S}}(1-\beta_\mathrm{RO})\right] + 4\beta_\mathrm{RO}\frac{\tau}{T_\mathrm{S}}\cos\left[\pi\frac{\tau}{T_\mathrm{S}}(1+\beta_\mathrm{RO})\right]}{\pi\frac{\tau}{T_\mathrm{S}}\left[1-\left(4\beta_\mathrm{RO}\frac{\tau}{T_\mathrm{S}}\right)^2\right]} & ,\text{otherwise} \end{cases}, \tag{3.53}$$

where $\beta_\mathrm{RO} \in (0, 1]$ is the *roll-off factor*. The exceptions for $\tau = 0$ and $\tau = \pm\frac{T_\mathrm{S}}{4\beta_\mathrm{RO}}$ can be shown using L'Hôpital's rule. Further if $\beta_\mathrm{RO} \to 0$ the equation is simplified to the sinc-pulse. In this case, the second line in the equation has to be ignored since $\tau = \pm\frac{T_\mathrm{S}}{4\beta_\mathrm{RO}}$ is undefined for $\beta_\mathrm{RO} = 0$. Another way to view the limit case $\beta_\mathrm{RO} \to 0$ is that in this case $\tau = \pm\frac{T_\mathrm{S}}{4\beta_\mathrm{RO}} \to \pm\infty$ and the expression in the second line evaluates to zero, thus correctly stating the limit-case for $\tau = \pm\infty$. Parameter β_RO has an influence on the pulses bandwidth which is given by

$$B_\mathrm{RRC} = \frac{1}{T_\mathrm{S}}\left(\beta_\mathrm{RO} + 1\right). \tag{3.54}$$

A small TBP is beneficial in regards to the systems SE, while a larger β_{RO} reduces inter symbol interference (ISI) e.g. in case of timing-jitter [35, Ch. 9.2.1]. For β_{RO} Eq. (3.53) is reduced to a sinc-pulse with bandwidth $B_{\mathrm{rrc}}(\beta_{\mathrm{RO}} = 0) = 1/T_{\mathrm{S}}$ while for $\beta_{\mathrm{RO}} = 1$ the occupied bandwidth is doubled. The additional bandwidth for $\beta_{\mathrm{RO}} > 0$ is called *excess bandwidth*. Two trains of normalized RRC-pulses for different values of β_{RO} are depicted in Fig. 3.12.

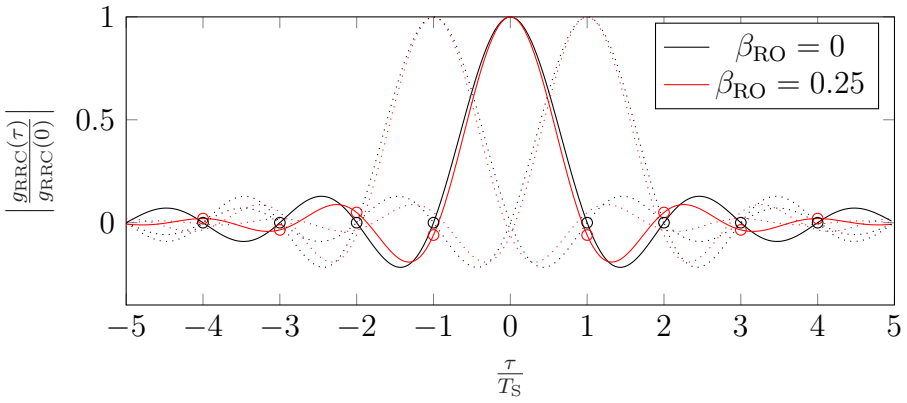


Figure 3.12.: RRC pulse-trains for two different roll-off factors β_{RO}.

For the RRC pulse with $\beta_{\mathrm{RO}} \neq 0$ the pulse is not longer zero at sampling times of neighboring pulses, resulting in deterministic ISI. However, if a *matched-filter* is used at the receiver, the Nyquist-criterion is fulfilled again, for all values $\beta_{\mathrm{RO}} > 0$.

The frequency spacing, given by f_k in Eq. (3.52), has to be chosen such that channels of width B_{RRC} do not overlap in frequency domain. Assuming, the same bandwidth is allocated for each user, a depiction of the frequency domain WDM signal is shown in Fig. 3.13. The additional bandwidth B_{G} is the *guard-band* and

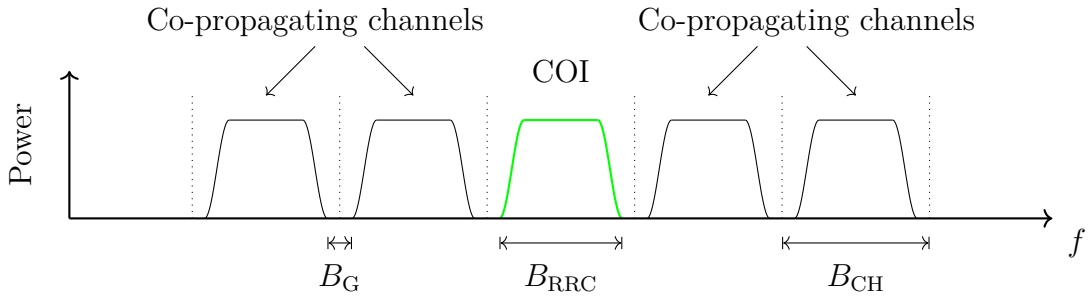


Figure 3.13.: WDM signal in frequency domain

is used to reduce nonlinear interference between channels during propagation [81] and

relax the requirements for e.g. receiver filters [82, Ch. 2.3]. The channel bandwidth $B_{\mathrm{CH}} = B_{\mathrm{RRC}} + B_{\mathrm{G}}$ is the total bandwidth one channel occupies. While in a real system, each channel in frequency domain is transmitting data, in simulations it is common to only measure the performance of the COI. The co-propagating channels are often just used to introduce nonlinear interference.

At the receiver, the transmitted data has to be retrieved from signal $r(\tau)$, which is the signal $s(\tau)$ after propagation over some channel. Detection is facilitated by matched-filtering, such that the received constellation symbol is obtained by [83]

$$a'_{nk} = \int_{-\infty}^{\infty} r(\tau) \cdot g^*_{\mathrm{TX}}(\tau - nT_{\mathrm{S}}) \cdot \mathrm{e}^{-\mathrm{j}2\pi f_k \tau} \mathrm{d}\tau, \tag{3.55}$$

where a'_{nk} is an estimate of the transmitted symbol a_{nk}. It can be seen from Eq. (3.52), that in order for Eq. (3.55) to be valid, the energy of $g_{\mathrm{TX}}(\tau)$ has to be normalized to $E_{g_{\mathrm{TX}}} = 1$. Matched filtering optimized the output SNR at the receiver and thus is used for detection in WDM systems over the course of this thesis.

3.5.2. Digital Signal Processing

If the transmitted signal $s(\tau)$ is propagating e.g. according to the channel governed by the NLSE, the signal is distorted due to linear and nonlinear effects. Even without considering ASE noise, the deterministic perturbations can severely restrict data-transmission by e.g. limiting the achievable transmission distance. Mitigation of linear effects in the physical domain can be done by e.g. introducing dispersion-managed links [51, Ch. 3.4.3] and there are also ways to combat nonlinear effects [84, Ch. 7.3]. However, DSP is enabling the mitigation of many linear and nonlinear effects at the receiver, while providing more flexibility since DSP components can be more easily replaced or reconfigured. In the following, two commonly used mitigation schemes will be reviewed.

Chromatic Dispersion Compensation

The chromatic dispersion compensation (CDC) algorithm is used to mitigate the effects of accumulated dispersion at the RX [50, Ch. 8.7.1]. CDC is treating the received signal as if it would have propagated through the linear channel described by Eq. (3.5), which is only valid if $P_{\mathrm{S}} \to 0$. Thus, the distortions, due to nonlinear effects can

not be compensated by this method. The received signal is modified according to

$$\hat{s}'(\omega) = \hat{r}(\omega) \cdot \mathrm{e}^{\mathrm{j}\omega^2\beta_2\mathcal{L}}, \tag{3.56}$$

where $\hat{x}(\omega)$ is the Fourier transform of $x(\tau)$ and $\hat{s}'(\omega)$ is the received signal after CDC, which should be similar to the transmitted signal $\hat{s}(\omega)$ for low signal powers and moderate additional ASE noise.

Digital Back-Propagation

Another commonly used DSP algorithm is digital back-propagation (DBP) [85]. This algorithm can mitigate dispersive and nonlinear effects jointly, by essentially performing a SSFM simulation of the inverse fiber channel. The linear and nonlinear parameters for the algorithm are given by

$$\hat{D}_{\mathrm{dbp}} = \frac{\alpha}{2} - \frac{\mathrm{j}\beta_2}{2}\frac{\partial^2}{\partial\tau^2}, \tag{3.57}$$

$$\hat{N}_{\mathrm{dbp}} = \mathrm{j}\gamma\left|r(\tau,\ell)\right|^2. \tag{3.58}$$

They can be used with each of the SSFM algorithms described in section 3.1.5. Note that, in the absence of noise and in the limit of $\Delta\ell \to 0$, this method can ideally compensate the distortions inflicted by the NLSE-channel, assuming a point-to-point scenario as described in section 3.5.3. In reality, often a rather low number of SSFM steps per span (N_{dbp}) suffice to significantly improve the transmission quality for moderate signal powers. Depending on the number of steps used, the complexity of DBP is rather high compared to other algorithms and might pose a bottleneck to data-rates or be cost-prohibitive for certain systems [86]. In recent years, there have been several publications, presenting reduced complexity algorithms for DBP incorporating e.g. recent advances in the field of deep neural networks [86], [87].

3.5.3. Lower Bounds on Achievable Information Rates

Using the schemes and algorithms presented above, systems for different transmission scenarios can be studied. In the following, two scenarios, namely *point-to-point transmission* and the *network case* will be studied in terms of their AIRs.

Point-to-Point Transmission

First, it is assumed that some fiber optic channel, modeled by the NLSE, is used exclusively by one TX/RX pair. In this case, the signals in each of the respective channels in the frequency domain, co-propagate for the whole transmission distance. In Fig. 3.14 the layout of such a system, assuming lumped amplification by EDFAs, is given. Since all WDM channels are generated by the same TX and co-propagate

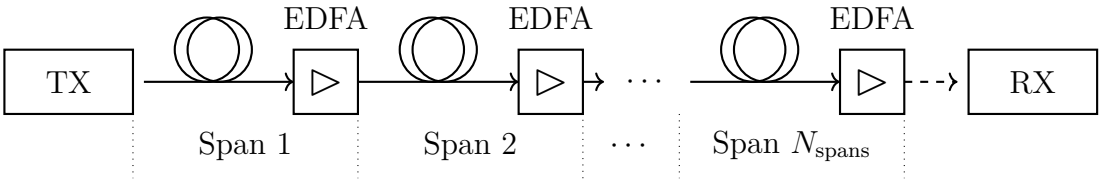


Figure 3.14.: Point-to-point optical communication scheme with lumped amplification

over the whole transmission distance, there is information about all WDM channels available at the RX. Thus, the complete received spectrum can be used by the DBP algorithm to mitigate nonlinear interactions between WDM channels. As previously noted, the complexity of this algorithm, which additionally scales with the number of WDM channels, poses a challenge for real-time implementations [88].

In Figure 3.15 the MI over SNR is given for a system with 16x16 MPSK modulation and a transmission distance of 20x100km. The noise power considered in the com-

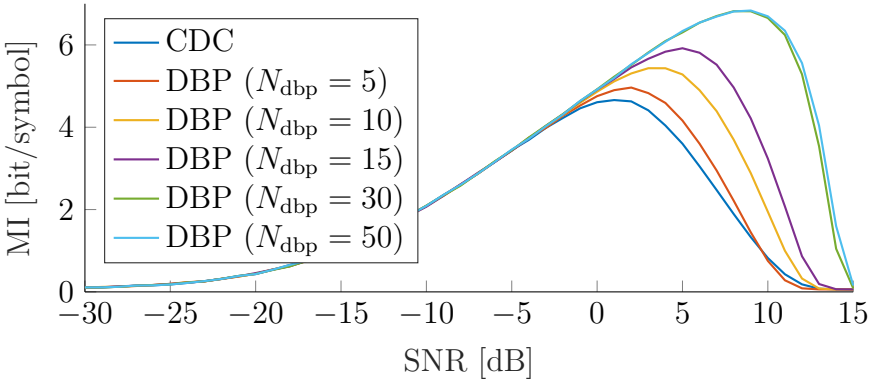


Figure 3.15.: 16x16 MPSK transmission over 20x100km (Point-to-point).

putation of the SNR is only including the ASE noise added by the amplifiers. More details on the specific system parameters can be found in Tab. C.1 in the appendix. It can be seen that, the peak of the curves shift towards higher values if the number

of steps for the DBP algorithm is increased. This effect saturates for high values, as can be seen for $N_{\text{dbp}} = 30$ and $N_{\text{dbp}} = 50$. The peak itself illustrates the fact that, in fiber optical communication systems using WDM, there is a *nonlinear limit* that results in an optimum launch-power for signals. This effect is not present in e.g. linear systems with additive noise. Note, that this nonlinear limit can not be overcome by algorithms such as DBP even if $N_{\text{dbp}} \to \infty$. While this would ideally mitigate linear and nonlinear perturbations in the noiseless case, the signal noise interaction can not be fully overcome even in this limit case. Thus, at least for the presented system, the described decrease in MI for high powers due to fiber nonlinearities poses a fundamental limit. Further, the peak shifts towards higher powers for more steps per span in the DBP algorithm. The optimum launch-power is also affected by other system parameters, e.g. by span length N_{span}. More results for other sets of parameters, can be found in the appendix in section D.1.

The Network Case

The network case differs from point-to-point transmission in the assumption, that the receiver has no information about the channels co-propagating with the COI. A depiction of such a network case scenario is given in Fig. 3.16. Here, several TX/RX

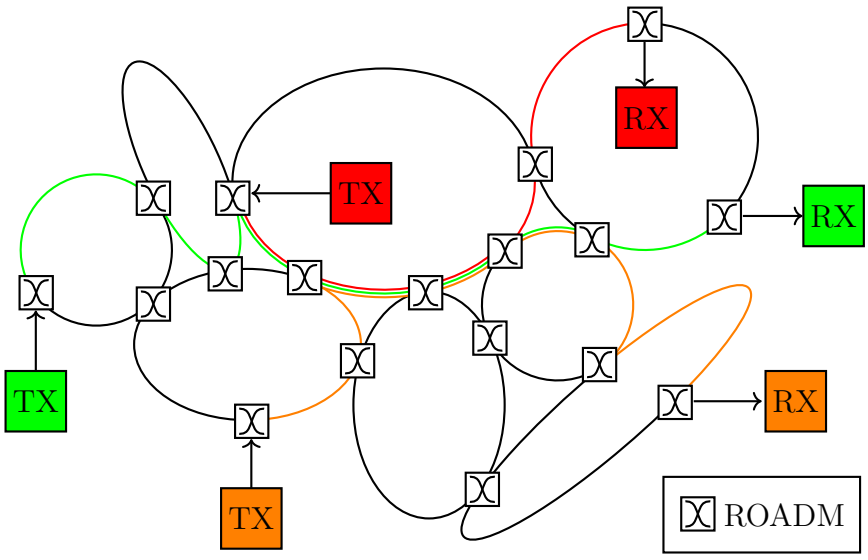


Figure 3.16.: Network scenario optical communication system

pairs share a network of fiber optic channels linked by reconfigurable optical add-drop multiplexers (ROADMs). The COI might thus co-propagate with other signals for a

certain distance, to then be redirected and co-propagate with another set of signals. As a result, the RX for the COI is assumed to have no information about other WDM-channels and can only utilize the COI for nonlinearity mitigation by e.g. DBP. The DBP algorithm can no longer mitigate the nonlinear interactions between the COI and potential co-propagating signals. In Fig. 3.17 the MI over SNR is given for a system identical to the point-to-point results shown in Fig. 3.15, using only the COI in the DBP algorithm. While the shown curves exhibit a similar behavior to the point-to-

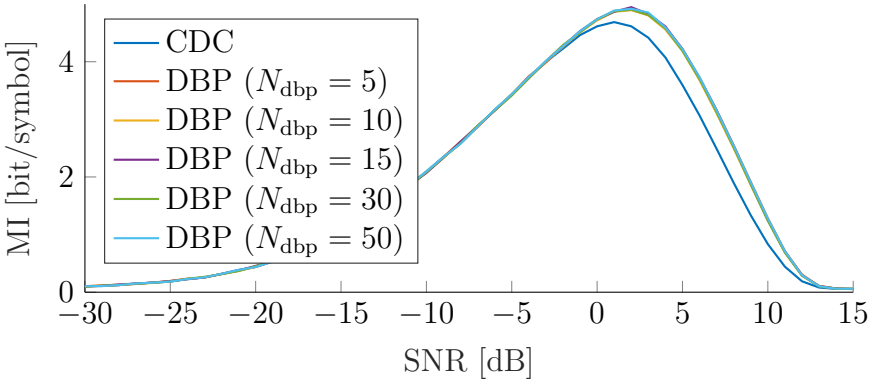


Figure 3.17.: 16x16 MPSK transmission over 20x100km (Network case).

point case, the achievable peak-values are lower in comparison. This is due to the fact, that nonlinearity is compensated less effectively, because the information about the co-propagating channels is not available. It can also be seen, that the gain achieved by using the DBP algorithm saturates for a lower number of steps per span. Similar to the point-to-point case, more results for the network case scenario with altered parameters, can be found in the appendix in section D.1.

4

The Nonlinear Fourier Transform for Single-Mode Fiber Channels

The main focus of this thesis is the implementation and study of NFT-aided communication systems. In reference to WDM systems, this type of transmission is also called nonlinear frequency division multiplexing (NFDM), indicating that data-streams can be multiplexed in the NFD accessed via the NFT and its inverse. The NFT originated as a tool to solve certain integrable partial differential equations [43]. The transform itself is channel-dependent and thus has to be explicitly developed for a fixed channel model. In this section, the NFT will be studied for the scalar, deterministic and lossless NLSE channel. Some results and insights are subsequently expanded for a multi-mode channel model in Chapter 5.

The transformations enable the treatment of signals in the NFD, in which the signals nonlinear frequency components do not interact with each other during propagation. Equalization of the intricate interplay between dispersion and nonlinearity in time domain is thus effectively dealt with, by inverting some multiplicative channel transfer function in the NFD. Thus, the relationship between the nonlinear channel and its NFT is analog to the relationship between the standard Fourier transform (FT) and linear convolutional channels. The concept of the NFT was called inverse scattering transform (IST) when it first was extensively studied in the 80s [89], [90]. While some authors still use the term IST up to this day [91], [92], the term NFT will be used over the course of this thesis, since it currently is widely adopted in literature and also

highlights the previously mentioned relation to the classic FT.

In the following, the mathematical basics for the NFT and its inverse will be introduced. Subsequently, some useful properties and closed form solutions for special signals will be given. A big portion of this section will then deal with the different algorithms that can be used to numerically implement the transformations. The chapter is concluded with several sections on special topics regarding the NFT and adjacent concepts and systems.

4.1. Basic Concept and Transformations

The scalar, deterministic and lossless NLSE channel is given by Eq. (3.1), neglecting the loss term on the right side of the equation. To simplify all further steps, Eq. (3.1) is normalized, such that it can be stated as [43]

$$\mathrm{j}q_z(t,z) = q_{tt}(t,z) + 2|q(t,z)|^2 q(t,z), \tag{4.1}$$

where $q(t,z)$, t, z are the normalized versions of $Q(\tau,\ell)$, τ, ℓ from Eq. (3.1). The z and t suffixes symbolize partial derivatives, such that e.g. $q_{tt}(t,z) = \partial^2 q(t,z)/\partial t^2$. For the sake of brevity, this shorthand notation will be used in the following derivations. The normalization step above is done using

$$T_0 = \sqrt{\frac{|\beta_2| L_0}{2}}, \tag{4.2}$$

$$P_0 = \sqrt{\frac{2}{\gamma L_0}}, \tag{4.3}$$

with

$$q(t,z) = \frac{Q(\tau,\ell)}{P_0}, \quad t = \frac{\tau}{T_0}, \quad z = \frac{\ell}{L_0}. \tag{4.4}$$

It can be seen, that one of the *normalization parameters* T_0, P_0 and L_0 can be chosen freely. Note, that in Eq. (4.1) it is assumed, that the channel is in the focusing regime ($\mathrm{sgn}(\beta_2) = -1$). This will be the prevalent assumption over the course of this thesis, since it allows for the generation of solitonic signals, which are discussed in detail in section 4.3.

4.1.1. Lax Pairs for Nonlinear Evolution Equations

The existence of a *Lax pair* [18] for the chosen channel model is a necessary basis for the development of the NFT. We start by stating [90, Ch. 1.2]

$$\mathrm{L}(z)v(\lambda,t,z) = \lambda v(\lambda,t,z), \tag{4.5}$$

$$v_z(\lambda,t,z) = \mathrm{M}(z)v(\lambda,t,z). \tag{4.6}$$

$\mathrm{L}(z)$ is some operator parameterized by z for which Eq. (4.5) is the spectral problem, λ are components of spectrum $\sigma(\mathrm{L}(z))$ and $v(\lambda,t,z)$ are the corresponding eigenfunctions. Operator $\mathrm{M}(z)$, also parameterized by z, governs the spatial evolution of the eigenfunctions of $\mathrm{L}(z)$ as stated in Eq. (4.6). Further, operator $\mathrm{L}(z)$ is similar to a multiplication operator Λ, satisfying the similarity transform [45, Ch. 2.7]

$$\mathrm{L}(z) = \mathrm{G}(z)\Lambda\mathrm{G}(z)^{-1}, \tag{4.7}$$

for some operator $\mathrm{G}(z)$

Taking the z-derivative of Eq. (4.7) and using Eq. (4.6) the *Lax equation* [18]

$$\mathrm{L}_z(z) = \mathrm{M}(z)\mathrm{L}(z) - \mathrm{L}(z)\mathrm{M}(z) = [\mathrm{M}(z), \mathrm{L}(z)]. \tag{4.8}$$

can be obtained. The operation $[\cdot,\cdot]$ is called the *commutator brackets*. According to [43, Lemma 1], $\sigma(\mathrm{L}(z))$ is independent of z if $\mathrm{L}(z)$ fulfills Eq. (4.8). Operator $\mathrm{L}(z)$ is thus *isospectral*. A pair of operators $\{\mathrm{L}(z), \mathrm{M}(z)\}$ is called a *Lax pair*, if it fulfills Eq. (4.8). A more detailed derivation of Eq. (4.8) can be found in appendix A.1, also yielding the relation $\mathrm{M}(z) = \mathrm{G}_z(z)\mathrm{G}^{-1}(z)$.

If operators $\{\mathrm{L}(z), \mathrm{M}(z)\}$ are now chosen in a suitable manner, Eq. (4.8) contains a nonlinear evolution equation of the form

$$q_z(t,z) = \mathrm{K}(q(t,z)). \tag{4.9}$$

Function $\mathrm{K}(\cdot)$ can generate terms, which are dependent on $q(t,z)$ and its temporal derivatives. The challenge is to find a pair of suitable operators for the evolution equation of interest. The Lax pairs for the nonlinear evolution equations studied in this thesis are given explicitly in sections 4.1.3 and 5.1 respectively. Note, that in some publications, including the initial paper by Lax [18], the role of temporal and spacial

parameters are switched. Nonetheless the definition shown above is adopted, since it is more prevalent in literature on optical communications.

4.1.2. The Zero-Curvature Condition

Using the Lax equation is one way to retrieve nonlinear evolution equations from a chosen Lax pair. Another relation that can be used is the *zero-curvature condition*. Starting from the temporal propagation equation for eigenfunctions of $\mathrm{L}(z)$

$$v_t(\lambda, t, z) = \mathrm{P}(z)v(\lambda, t, z) \tag{4.10}$$

with some operator $\mathrm{P}(z)$ and Eq. (4.6), the zero curvature condition can be obtained by equating the spatial derivative of Eq. (4.10) with the temporal derivative of Eq. (4.6) [90, Ch. 1.2]. The zero curvature condition is

$$\mathrm{P}_z - \mathrm{M}_t + [\mathrm{P}, \mathrm{M}] = 0. \tag{4.11}$$

In the following, the z-dependencies of operators are sometimes omitted for the sake of brevity. If the operators P, M are chosen correctly, the zero-curvature condition once again contains a nonlinear evolution equation e.g. the NLSE [93, Ch. 1.2]. More importantly, Eq. (4.10) is a linear propagation equation for the eigenfunctions, that is important in the development of the NFT in later sections.

As will be seen in the next section, the operators $\mathbf{L}(z)$ and $\mathbf{P}(z)$ for the NLSE from Eqs. (4.14) and (4.15) can be related by straightforward algebraic manipulation of Eq. (4.5). There also is the more general relation [43]

$$\mathrm{P} = \Sigma(\mathrm{L} - \lambda\mathrm{I}) + \mathrm{DI}, \tag{4.12}$$

where Σ is some invertible operator and $\mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}t}$. It can be obtained, by rearranging and equating Eqs. (4.5) and (4.10). More details on the derivations of Eqs. (4.11) and (4.12) can be found in section A.2 in the appendix.

4.1.3. The Zakharov-Shabat System

For many equations that have practical significance the operator $\mathbf{L}(z)$ takes on the form of [15], [43], [89]

$$\mathbf{L}(z) = \mathrm{j} \begin{bmatrix} \mathrm{D} & -r(t,z) \\ s(t,z) & -\mathrm{D} \end{bmatrix} \tag{4.13}$$

For the NLSE $r(t) = -q(t)$ and $s(t) = -q(t)^*$ yields operator [43]

$$\mathbf{L}(z) = \mathrm{j} \begin{bmatrix} \mathrm{D} & -q(t,z) \\ -q^*(t,z) & -\mathrm{D} \end{bmatrix}. \tag{4.14}$$

For this particular case, the operator $\mathbf{P}(z)$ can be found by rearranging Eq. (4.5), after setting the operator to Eq. (4.14). It has the form

$$\mathbf{P}(z) = \begin{bmatrix} -\mathrm{j}\lambda & q(t,z) \\ -q^*(t,z) & \mathrm{j}\lambda \end{bmatrix}. \tag{4.15}$$

If the definitions above are used in Eq. (4.11) and the result is compared to Eq. (4.1), the spatial propagation operator for the eigenvectors is found to be

$$\mathbf{M}(z) = \begin{bmatrix} 2\mathrm{j}\lambda^2 - \mathrm{j}|q(t,z)|^2 & -2\lambda q(t,z) - \mathrm{j}q_t(t,z) \\ 2\lambda q^*(t,z) - \mathrm{j}q_t^*(t,z) & -2\mathrm{j}\lambda^2 + \mathrm{j}|q(t,z)|^2 \end{bmatrix}. \tag{4.16}$$

Utilizing Eq. (4.10) and (4.15) the Zakharov-Shabat (ZS) system is given by [15]

$$\boldsymbol{v}_t(\lambda,t,z) = \mathbf{P}(z)\boldsymbol{v}(\lambda,t,z) = \begin{bmatrix} -\mathrm{j}\lambda & q(t,z) \\ -q^*(t,z) & \mathrm{j}\lambda \end{bmatrix} \begin{bmatrix} v_1(\lambda,t,z) \\ v_2(\lambda,t,z) \end{bmatrix}, \tag{4.17}$$

where $v_i(\lambda,t,z)$ are the entries of $\boldsymbol{v}(\lambda,t,z)$. Eq. (4.17) is the basis for the development of the NFT for the scalar NLSE channel.

4.1.4. Boundary Conditions and the Wronskian

The NFT transformation is always done for some specific spatial value z. Thus, the z-dependence of quantities is omitted in the following. To find a set of boundary conditions for the ZS system, some assumptions on $q(t)$ have to be made. Function

$q(t)$ here plays the role of the signal pulse transmitted over the fiber optic channel. Thus, restricting the pulses energy to finite values and requiring the pulse to decay sufficiently fast for $t \to \infty$ is reasonable. The imposed restrictions are [43], [90]

$$\int_{-\infty}^{\infty} |q(t)|^2 < \infty \tag{4.18}$$

$$q(t) \to 0 \text{ for } |t| \to \infty. \tag{4.19}$$

Note, that it is also possible to derive a NFT for periodic boundary conditions as detailed in [94], [95]. However, over the course of this thesis the conditions from Eqs. (4.18) and (4.19) apply. This results in Eq. (4.17) having form

$$\lim_{|t|\to\infty} \boldsymbol{v}_t(\lambda, t) = \begin{bmatrix} -\mathrm{j}\lambda & 0 \\ 0 & \mathrm{j}\lambda \end{bmatrix} \boldsymbol{v}(\lambda, t), \tag{4.20}$$

for $|t| \to \infty$ and thus its closed form solution at this limit is

$$\lim_{|t|\to\infty} \boldsymbol{v}(\lambda, t) = \begin{bmatrix} c_1 \cdot \mathrm{e}^{-\mathrm{j}\lambda t} \\ c_2 \cdot \mathrm{e}^{\mathrm{j}\lambda t} \end{bmatrix}. \tag{4.21}$$

The components λ of spectrum $\sigma(\mathbf{L})$ can attain any value in the complex plane and thus have the form $\lambda = \xi + \mathrm{j}\eta$ with $\xi, \eta \in \mathbb{R}$. For now it is assumed that $\lambda \in \mathbb{R}$ is on the real line and thus, the boundary conditions can be defined as

$$\boldsymbol{\phi}(\lambda, t) \to \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathrm{e}^{-\mathrm{j}\lambda t} \qquad \text{for } t \to -\infty, \tag{4.22}$$

$$\bar{\boldsymbol{\phi}}(\lambda, t) \to \begin{bmatrix} 0 \\ -1 \end{bmatrix} \mathrm{e}^{\mathrm{j}\lambda t} \qquad \text{for } t \to -\infty, \tag{4.23}$$

$$\boldsymbol{\psi}(\lambda, t) \to \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathrm{e}^{\mathrm{j}\lambda t} \qquad \text{for } t \to \infty, \tag{4.24}$$

$$\bar{\boldsymbol{\psi}}(\lambda, t) \to \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathrm{e}^{-\mathrm{j}\lambda t} \qquad \text{for } t \to \infty, \tag{4.25}$$

where $\{\boldsymbol{\phi}(\lambda, t), \bar{\boldsymbol{\phi}}(\lambda, t), \boldsymbol{\psi}(\lambda, t), \bar{\boldsymbol{\psi}}(\lambda, t)\}$ are called the *Jost solutions* of the ZS system. Note that in this step different boundary conditions can be used. This would result

in slight differences in the transformations, but would not impede the functionality of the transformation itself. Since the review of the basics of the NFT is presented along the lines of [90], the boundary conditions from the respective source are adopted.

The *Wronskian* $\mathrm{W}(\cdot,\cdot)$ of two functions $\boldsymbol{v}(\lambda,t) = [v_1(\lambda,t), v_2(\lambda,t)]^{\mathrm{T}}$ and $\boldsymbol{w}(\lambda,t) = [w_1(\lambda,t), w_2(\lambda,t)]^{\mathrm{T}}$ is defined as

$$\mathrm{W}(\boldsymbol{v}(\lambda,t), \boldsymbol{w}(\lambda,t)) = v_1(\lambda,t)w_2(\lambda,t) - v_2(\lambda,t)w_1(\lambda,t) \tag{4.26}$$

It can be shown that, if $\boldsymbol{v}(\lambda,t)$ and $\boldsymbol{w}(\lambda,t)$ are solutions of (4.17), the Wronskian is independent of time and thus [90]

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{W}(\boldsymbol{v}(\lambda,t), \boldsymbol{w}(\lambda,t)) = 0. \tag{4.27}$$

This can be shown, by first showing $\mathrm{dW}(\boldsymbol{v}(\lambda,t), \boldsymbol{w}(\lambda,t))/\mathrm{d}t = \mathrm{W}(\boldsymbol{v}_t(\lambda,t), \boldsymbol{w}(\lambda,t)) + \mathrm{W}(\boldsymbol{v}(\lambda,t), \boldsymbol{w}_t(\lambda,t))$ and then using Eq. (4.17) with this result. A detailed proof is given in [43, Lemma 2]. Looking at Eq. (4.26), one can also see that $\mathrm{W}(\boldsymbol{v}(\lambda,t), \boldsymbol{w}(\lambda,t)) \neq 0$ is a criterion for linear independence of the two solutions.

4.1.5. Nonlinear Fourier Coefficients in the Complex Plane

It can be seen from Eqs. (4.22) to (4.25) that, $\mathrm{W}(\boldsymbol{\phi}(\lambda,t), \bar{\boldsymbol{\phi}}(\lambda,t)) = -1$ as well as $\mathrm{W}(\boldsymbol{\psi}(\lambda,t), \bar{\boldsymbol{\psi}}(\lambda,t)) = -1$. Thus $\{\boldsymbol{\phi}(\lambda,t), \bar{\boldsymbol{\phi}}(\lambda,t)\}$ and $\{\boldsymbol{\psi}(\lambda,t), \bar{\boldsymbol{\psi}}(\lambda,t)\}$ form linearly independent sets and can be used in the projection equations

$$\boldsymbol{\phi}(\lambda,t) = a(\lambda)\bar{\boldsymbol{\psi}}(\lambda,t) + b(\lambda)\boldsymbol{\psi}(\lambda,t) \tag{4.28}$$

$$\bar{\boldsymbol{\phi}}(\lambda,t) = -\bar{a}(\lambda)\boldsymbol{\psi}(\lambda,t) + \bar{b}(\lambda)\bar{\boldsymbol{\psi}}(\lambda,t) \tag{4.29}$$

to define the *scattering coefficients* $\{a(\lambda), b(\lambda), \bar{a}(\lambda), \bar{b}(\lambda)\}$. From the property given in Eq. (4.27) and

$$\mathrm{W}(\boldsymbol{\phi}(\lambda,t), \boldsymbol{\psi}(\lambda,t)) = a(\lambda) \tag{4.30}$$

$$\mathrm{W}(\bar{\boldsymbol{\phi}}(\lambda,t), \bar{\boldsymbol{\psi}}(\lambda,t)) = \bar{a}(\lambda) \tag{4.31}$$

$$\mathrm{W}(\boldsymbol{\phi}(\lambda,t), \bar{\boldsymbol{\psi}}(\lambda,t)) = -b(\lambda) \tag{4.32}$$

$$\mathrm{W}(\bar{\boldsymbol{\phi}}(\lambda,t), \boldsymbol{\psi}(\lambda,t)) = \bar{b}(\lambda) \tag{4.33}$$

one can see that the scattering coefficients are time independent. Thus, the time instance to compute the scattering coefficients can be chosen freely, as long as all Jost solutions are available for the respective value of t. It is convenient to choose $t \to \infty$, since $\boldsymbol{\psi}(\lambda, t)$ and $\bar{\boldsymbol{\psi}}(\lambda, t)$ are given by Eqs. (4.24) and (4.25). In most cases, it is sufficient to only obtain $\{a(\lambda), b(\lambda)\}$, using Eq. (4.28). Thus, the boundary condition in Eq. (4.22) has to be propagated to $t \to \infty$. Coefficients $\{a(\lambda),\, b(\lambda)\}$ are called the *nonlinear Fourier coefficients*.

Eqs. (4.28) and (4.29) are well defined as long as $\lambda \in \mathbb{R}$. If this assumption is relaxed, assuming that now $\lambda = \xi + \mathrm{j}\eta$ with $\eta > 0$ can be any point in the positive complex halfplane $\mathbb{C}^+$, it can be shown that $\mathrm{e}^{\mathrm{j}\lambda t}\boldsymbol{\phi}(\lambda, t)$ and $\mathrm{e}^{-\mathrm{j}\lambda t}\boldsymbol{\psi}(\lambda, t)$ are analytic functions of $\lambda \in \mathbb{C}^+$ [89, Ch. IV.A]. Thus, $a(\lambda)$ is analytic in $\mathbb{C}^+$ as well (see Eq. (4.30)). It can be shown, that $\bar{a}(\lambda)$ is analytic in $\mathbb{C}^-$. It is not necessary for $b(\lambda)$ and $\bar{b}(\lambda)$ to be analytic anywhere, but it can be shown that if function $q(t)$ decays faster than $C_1\mathrm{e}^{-2C_2|t|}$ for some constants C_1 and C_2, then $b(\lambda)$ and $\bar{b}(\lambda)$ are analytic for values $\lambda = \xi + \mathrm{j}\eta$ satisfying $C_2 > \eta > -C_2$ [90, Ch. 1.3]. The treatment of the given functions analyticity is rather involved and many subtleties do not come up in the practical use of the derived transformations. Thus, further necessary results regarding the nonlinear Fourier coefficients in $\mathbb{C}$ will be given in the following, without stating details on their derivations. The interested reader is referred to the in-depth treatments presented in respective literature cited below.

Since $a(\lambda)$ is analytic in $\mathbb{C}^+$, any value λ_k with $1 \leq k \leq K$ for which $a(\lambda_k) = 0$ must be an isolated point [96, Theorem 8.14] and K is in general finite [89, Ch. IV.A]. For a pulse $q(t)$ with sufficient energy there thus exist finitely many *discrete eigenvalues* λ_k for which Eq. (4.28) is bounded and simplified to

$$\boldsymbol{\phi}(\lambda_k, t) = b(\lambda_k)\boldsymbol{\psi}(\lambda_k, t). \tag{4.34}$$

A similar observation can be made for $\bar{a}(\lambda)$ and Eq. (4.29) in the negative complex half-plane $\mathbb{C}^-$. Further, for the ZS system, for any discrete eigenvalue $\lambda_k \in \mathbb{C}^+$ there exists a discrete eigenvalue $\lambda_k{}^* \in \mathbb{C}^-$ [43, Lemma 2]. Thus, it is sufficient to consider the real line $\lambda \in \mathbb{R}$ and the discrete eigenvalues in the positive complex half-plane $\lambda_k \in \mathbb{C}^+$. Over the course of the thesis, it will also be assumed that all zeros of $a(\lambda)$ in the positive complex half-plane are simple. This can be justified, since the signals used in later transmission scenarios are constructed, such that they always fulfill this assumption. For more information on zeros of higher multiplicity, the interested reader

is referred to [90], [97].

4.1.6. The Nonlinear Fourier Spectra

Looking at Eqs. (4.28) and (4.29), it can be seen that, if e.g. the equations are evaluated for $t \to \infty$, the boundary conditions for $t \to -\infty$ have to be propagated towards infinity. This makes them interact with the pulse $q(t)$ at finite values, which alters the resulting nonlinear Fourier coefficients $\{a(\lambda), b(\lambda)\}$. While not apparent immediately, it can be shown that the nonlinear Fourier coefficients contain "complete" information about the pulse $q(t)$ and knowledge of $\{a(\lambda), b(\lambda)\}$ is sufficient to recover pulse $q(t)$ [43, Sec. VII]. Further, to fully reconstruct $q(t)$ it even suffices to just have knowledge about the following two quantities. The first quantity is the *continuous nonlinear spectrum*

$$q_{\mathrm{c}}(\lambda) = \frac{b(\lambda)}{a(\lambda)}. \tag{4.35}$$

which is defined on the whole real line $\lambda \in \mathbb{R}$.

The second part consists of the *discrete eigenvalues* λ_k in the positive complex half-plane $\mathbb{C}^+$. The corresponding *discrete spectral amplitudes* are defined as

$$q_{\mathrm{d}}(\lambda_k) = \frac{b(\lambda_k)}{a_\lambda(\lambda_k)}, \tag{4.36}$$

for a set of discrete eigenvalues λ_k for $1 \leq k \leq K < \infty$, where $a_\lambda(\lambda_k) = \frac{\mathrm{d}a(\lambda)}{\mathrm{d}\lambda}\Big|_{\lambda=\lambda_k}$. Since only simple zeros of $a(\lambda)$ are considered and it is assumed that the pulse $q(t)$ decays sufficiently fast, such that $b(\lambda)$ is analytic for all λ_k with $1 \leq k \leq K$, the equation for the discrete spectral amplitudes in Eq. (4.36) can directly be derived from Eq. (4.35) and the relation for the residue $\operatorname*{Res}_{\lambda=\lambda_k}(b(\lambda)/a(\lambda)) = b(\lambda)/a_\lambda(\lambda)$ [96, Proposition 9.14].

Note, that even though λ is part of the spectrum of $\mathbf{L}$, the quantities $q_{\mathrm{c}}(\lambda)$ associated with λ are called the continuous nonlinear spectrum. Similarly, the *discrete nonlinear spectrum*, consists of the discrete eigenvalues λ_k and either the discrete spectral amplitudes $q_{\mathrm{d}}(\lambda_k)$ or the values $b(\lambda_k)$, depending on the exact definition used. While not immediately intuitive, this nomenclature is widely used in publications in the field and thus adopted in this thesis.

The continuous nonlinear spectrum represents the *radiative* part of the function $q(t)$, which does not show solitonic behavior during propagation. The discrete nonlinear

spectrum corresponds to the solitonic part of the pulse $q(t)$. This solitonic component, while it may periodically change shape during propagation, will not become infinitely broad even if the signal is transmitted for an infinite distance. Rather, the pulses energy will remain temporally confined on a finite temporal support. A more in-depth look at this behavior is taken in section 4.3.

4.1.7. The Lax Convolution Channel

Similarly to how, for linear systems, the influence of the channel on the signal at the receiver ($z = \mathcal{L}$) can be described by a convolution of the signal at the transmitter ($z = 0$) with the channels impulse response $h(t)$, the *Lax convolution channel* [43, Sec. III] can be used to describe communication over an *integrable* [98] channel, such as the fiber-optic communication channel, modeled by equation (4.1). Using a similar notation, the received signal can be described by *Lax convolution*

$$q(t, z = \mathcal{L}) = q(t, z = 0) * (\mathbf{L}(z), \mathbf{M}(z); \mathcal{L}) \tag{4.37}$$

where $(\mathbf{L}(z), \mathbf{M}(z); \mathcal{L})$ denotes the integrable system by its respective Lax pair and propagation distance. Another way to signify this is given in Fig. 4.1. An elaborate

$$q(t, z = 0) \longrightarrow \boxed{\mathbf{L}_z(z) = [\mathbf{M}(z), \mathbf{L}(z)]} \longrightarrow q(t, z = \mathcal{L})$$

Figure 4.1.: Integrable system defined by a Lax convolution

discussion on what constitutes an integrable communication channel can be found in [99]. Over the course of this thesis, we will assume that as long as the noise introduced into the deterministic NLSE channel is small enough, we can still assume integrability of the optical fiber channel [43].

4.1.8. The Nonlinear Fourier Transform

The NFT is a transformation from the time domain to the nonlinear frequency domain, yielding the nonlinear Fourier spectrum, which is often separated into two disjoint parts, namely the continuous ($q_c(\lambda)$) and the discrete nonlinear spectrum

($\{q_\mathrm{d}(\lambda_k), \lambda_k\}$). The transformation is defined as

$$\mathrm{NFT}(q(t)) = \{q_\mathrm{c}(\lambda), q_\mathrm{d}(\lambda_k), \lambda_k\} \text{ for } \lambda \in \mathbb{R}, \lambda_k \in \mathbb{C}^+, 1 \leq k \leq K. \tag{4.38}$$

Similarly, the inverse nonlinear Fourier transform (INFT) from the nonlinear frequency domain to the time domain is defined as

$$\mathrm{INFT}(\{q_\mathrm{c}(\lambda), q_\mathrm{d}(\lambda_k), \lambda_k\}) = q(t). \tag{4.39}$$

For certain pulses, only certain parts of the nonlinear spectrum are present. If a pulse does not have any non-zero components in the continuous spectrum ($q_\mathrm{c}(\lambda) = 0 \,\forall\, \lambda \in \mathbb{R}$) but has at least one discrete eigenvalue, the resulting time-domain pulse is called *discrete spectrum pulse* and will always be a solitonic solution of Eq. (4.1). On the other hand, if a pulse does not have a discrete spectrum, i.e. there are no $\lambda_k \in \mathbb{C}^+$ for which $a(\lambda_k) = 0$, but at least some values of $q_\mathrm{c}(\lambda)$ on the real line are non-zero, the pulse is called *continuous spectrum pulse*, being a purely radiative solution of the NLSE. In the most general case, where discrete eigenvalues exist in the upper-half complex plane and $q_\mathrm{c}(\lambda)$ has non-zero components, the pulse consists of radiative and solitonic components. If no discrete eigenvalues exist and the continuous spectrum is zero everywhere on the real line, the pulse will be the zero pulse ($q(t) = 0 \;\forall\; t$). This last relation will also become obvious from the energy considerations presented in section 4.2.3. While the composition of a pulses nonlinear spectrum will influence the choice of algorithms to use, the notation (NFT($\cdot$), INFT($\{\cdot, \cdot, \cdot\}$)) above is used for all described cases.

4.2. Properties of the Nonlinear Spectra

Some fundamental properties of the NFT for the NLSE were already given in section 4.1. In the following, some other useful properties are discussed and presented. A big portion of the shown material is directly taken from [43] and [90]. The reader is thus referred to this sources for most of the proofs and a more in-depth treatment.

4.2.1. Phase, Time and Frequency Shifts

Similar to the standard FT, relations for the influence of phase, time and frequency shifts on the nonlinear Fourier spectrum can be obtained.

Constant Phase Change

This property relates the effect of a constant phase change in time domain to the nonlinear spectra. For some constant phase φ and continuous and discrete nonlinear spectra $\{q_\mathrm{c}(\lambda), q_\mathrm{d}(\lambda_k), \lambda_k\}$ related to $q(t)$, it can be shown that [43, Sec. 4.D, App. B]

$$\mathrm{NFT}\left(\mathrm{e}^{-\mathrm{j}\varphi}q(t)\right) = \left\{\mathrm{e}^{\mathrm{j}\varphi}q_\mathrm{c}(\lambda), \mathrm{e}^{\mathrm{j}\varphi}q_\mathrm{d}(\lambda_k), \lambda_k\right\}. \tag{4.40}$$

Note that the nonlinear spectral amplitudes $\{q_\mathrm{c}(\lambda), q_\mathrm{d}(\lambda_k)\}$ are affected by an inverse phase shift and the discrete eigenvalues λ_k are not altered by the phase change.

Time Shift

If $q(t)$ is shifted in time by some fixed value t_0, the corresponding nonlinear spectra are modified according to [43, Sec. 4.D, App. B]

$$\mathrm{NFT}\left(q(t-t_0)\right) = \left\{\mathrm{e}^{-2\mathrm{j}\lambda t_0}q_\mathrm{c}(\lambda), \mathrm{e}^{-2\mathrm{j}\lambda_k t_0}q_\mathrm{d}(\lambda_k), \lambda_k\right\}. \tag{4.41}$$

Frequency Shift

Similarly, if the time domain signal $q(t)$ is shifted by some fixed value ω_0 in the linear frequency domain, the corresponding nonlinear spectra are changing according to [43, Sec. 4.D, App. B]

$$\mathrm{NFT}\left(q(t)\mathrm{e}^{-2\mathrm{j}\omega_0 t}\right) = \left\{q_\mathrm{c}(\lambda-\omega_0), q_\mathrm{d}(\lambda_k-\omega_0), \lambda_k-\omega_0\right\}. \tag{4.42}$$

4.2.2. The Layer-Peeling Property

The *layer-peeling* property can be used to compute joint nonlinear Fourier coefficients for temporally non-overlapping functions. It is most prominently utilized in the layer-peeling algorithm presented in section 4.4.1. If two time domain signals $q_1(t)$, $q_2(t)$ can be represented by the nonlinear Fourier coefficients $\{a_1(\lambda), b_1(\lambda)\}$ and $\{a_2(\lambda), b_2(\lambda)\}$ and have disjoint supports $[t_1, t_2]$, $[t_2, t_3]$ respectively, the joint signal $q(t) = q_1(t)+q_2(t)$ is represented by nonlinear Fourier coefficients $\{a(\lambda), b(\lambda)\}$, which can be computed

according to [27]

$$\begin{aligned}\{a(\lambda), b(\lambda)\} =& \{a_1(\lambda), b_1(\lambda)\} \circ \{a_2(\lambda), b_2(\lambda)\} \\ =& \{a_1(\lambda)a_2(\lambda) - b_1(\lambda)b_2^*(\lambda^*), a_1(\lambda)b_2(\lambda) + b_1(\lambda)a_2^*(\lambda^*)\}. \end{aligned} \tag{4.43}$$

4.2.3. The Trace Formula

The trace formula [43, App. B] gives conserved quantities for the time-domain signal $q(t)$. The first three conserved quantities are the energy, momentum and Hamiltonian and are given by [27, Sec. 4.A]

$$C^{(1)} = \int_{-\infty}^{\infty} |q(t)|^2 \, \mathrm{d}t, \tag{4.44}$$

$$C^{(2)} = \frac{1}{2\mathrm{j}} \int_{-\infty}^{\infty} q(t) \frac{\mathrm{d}q^*(t)}{\mathrm{d}t} \mathrm{d}t, \tag{4.45}$$

$$C^{(3)} = \int_{-\infty}^{\infty} -\frac{1}{4} \left(|q(t)|^4 - \left| \frac{\mathrm{d}q(t)}{\mathrm{d}t} \right|^2 \right) \mathrm{d}t. \tag{4.46}$$

These quantities can be related to the following quantities, derived for the continuous and discrete nonlinear spectrum, according to $C^{(i)} = C^{(i)}_{\mathrm{cont}} + C^{(i)}_{\mathrm{disc}}$ using [100], [90, Ch. 1.6]

$$C^{(i)}_{\mathrm{cont}} = \frac{1}{\pi} \int_{-\infty}^{\infty} \lambda^{i-1} \log(1 + |q_{\mathrm{c}}(\lambda)|^2) \mathrm{d}\lambda, \tag{4.47}$$

$$C^{(i)}_{\mathrm{disc}} = \frac{4}{i} \sum_{k=1}^{K} \Im\{\lambda_k^i\}. \tag{4.48}$$

In many cases, the energy relation is the most useful and can be explicitly stated as

$$\int_{-\infty}^{\infty} |q(t)|^2 \, \mathrm{d}t = \underbrace{\frac{1}{\pi} \int_{-\infty}^{\infty} \lambda^{i-1} \log(1 + |q_{\mathrm{c}}(\lambda)|^2) \mathrm{d}\lambda}_{\text{Cont. Spectrum}} + \underbrace{\frac{4}{i} \sum_{k=1}^{K} \Im\{\lambda_k^i\}}_{\text{Disc. Spectrum}}, \tag{4.49}$$

where Eqs. (4.47) and (4.48) were used with $i = 1$. The above equation is frequently used to check if the energy of the nonlinear spectrum, generated by some NFT algorithm, approximately matches the energy of the time-domain pulse. This could e.g.

indicate the success or failure of a numerical transformation and is especially useful for the computation of discrete eigenvalues, as will be discussed in section 4.4.2. A more in-depth treatment of conserved quantities of the time-domain signal and the respective nonlinear spectra is given in [90, Ch. 1.6].

4.2.4. Propagation of Nonlinear Spectra in the NLSE Channel

Time domain signal $q(t)$, propagating in the NLSE channel is subject to an intricate interplay between dispersion and nonlinearity. One of the primary advantages, of treating the propagating signal through its nonlinear spectrum, is that the influence of the NLSE channel on $q(t)$, can be captured by the simple relations

$$q_{\mathrm{c}}(\lambda, z) = \mathrm{e}^{-4\mathrm{j}\lambda^2 z} q_{\mathrm{c}}(\lambda, 0), \tag{4.50}$$

$$q_{\mathrm{d}}(\lambda_k, z) = \mathrm{e}^{-4\mathrm{j}\lambda_k^2 z} q_{\mathrm{d}}(\lambda_k, 0), \tag{4.51}$$

$$b(\lambda_k, z) = \mathrm{e}^{-4\mathrm{j}\lambda_k^2 z} b(\lambda_k, 0). \tag{4.52}$$

The relation between time domain and NFD is depicted in Fig. 4.2. It can be seen,

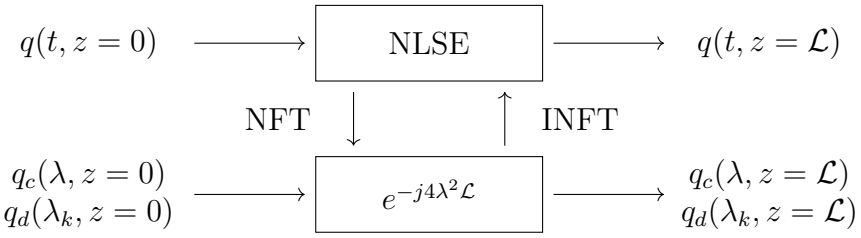


Figure 4.2.: NLSE channel in time and nonlinear frequency domain.

that the nonlinear spectral components for each nonlinear frequency λ propagate independently, while the nonlinear frequencies, including the discrete eigenvalues λ_k, are constant in z, as already discussed in section 4.1.1. This enables the joint equalization of nonlinearity and dispersion, by inverting the multiplicative channel response in the NFD. This compensation scheme exhibits very low complexity, as long as it is assumed that the nonlinear spectra are readily available, due to e.g. effective transformation algorithms. While this thesis does not have an explicit focus on the comparison between e.g. WDM and NFDM schemes in terms of their complexity, it is still important to note, that in order to outperform traditional WDM the SE of an NFDM system has to surpass WDM, while the whole transmission scheme, including the transforma-

tions, exhibits comparable or smaller complexity than the respective WDM system. In recent years, some low complexity algorithms for the NFT and their inverse were published [30], [31], [101], [102].

4.3. Special Solutions

In most cases, it is not feasible to compute the NFT or its inverse analytically. For these general cases, several numerical algorithms exist, that will be discussed in detail in section 4.4. However, there are analytic expressions for the nonlinear spectrum of some specific signals. In the following, some well known examples will be presented, as the results are useful either to find numerically favorable algorithms for the transformations, or to serve as a way to evaluate the accuracy of existing transformations in later sections.

4.3.1. The Rectangular Pulse

For the NFT of the rectangular pulse

$$q_{\mathrm{rec}}(t) = \begin{cases} A, & t \in [t_1, t_2] \\ 0, & \text{else} \end{cases}, \tag{4.53}$$

where $T_{\mathrm{rec}} = t_2 - t_1$ is the pulse-width and A is the pulses amplitude, some closed-form expressions can be found. For some fixed distance z, the operator $\mathbf{P}(z) = \boldsymbol{P}$ is constant on interval $[t_1, t_2]$. Thus, the vector $\boldsymbol{\phi}(\lambda, t)$ can be computed in closed form using Eq. (4.17). This results in

$$\boldsymbol{\phi}(t, \lambda) = \mathrm{e}^{(t-t_1)\boldsymbol{P}} \boldsymbol{\phi}(t_1, \lambda), \ t \in [t_1, t_2], \tag{4.54}$$

where the boundary condition for t_1 is $\boldsymbol{\phi}(t_1, \lambda) = [1, 0]^{\mathrm{T}} \mathrm{e}^{-\mathrm{j}\lambda t_1}$. Outside of interval $[t_1, t_2]$, the pulse is zero and thus $\boldsymbol{\psi}(\lambda, t_2) = [0, 1]^{\mathrm{T}} \mathrm{e}^{\mathrm{j}\lambda t_2}$. Vector $\boldsymbol{\phi}(t_2, \lambda) = \mathrm{e}^{T_{\mathrm{rec}}\boldsymbol{P}} \boldsymbol{\phi}(t_1, \lambda)$, where constant $\boldsymbol{P}$ is a 2x2 matrix and thus the matrix exponential is given by a power-

series. The matrix-exponential can be computed in closed form, yielding [43, Sec. 4.C]

$$\begin{aligned} \mathrm{e}^{T_{\mathrm{rec}}\boldsymbol{P}} =& \exp\left(\begin{bmatrix} -\mathrm{j}\lambda & A \\ -A^* & \mathrm{j}\lambda \end{bmatrix} T_{\mathrm{rec}}\right) \\ =& \begin{bmatrix} \cos(\Delta T_{\mathrm{rec}}) - \mathrm{j}\frac{\lambda}{\Delta}\sin(\Delta T_{\mathrm{rec}}) & \frac{A}{\Delta}\sin(\Delta T_{\mathrm{rec}}) \\ \frac{-A^*}{\Delta}\sin(\Delta T_{\mathrm{rec}}) & \cos(\Delta T_{\mathrm{rec}}) + \mathrm{j}\frac{\lambda}{\Delta}\sin(\Delta T_{\mathrm{rec}}) \end{bmatrix}, \end{aligned} \tag{4.55}$$

with $\Delta = \sqrt{\lambda^2 + |A|^2}$. Using the results above and Eq. (4.28), the nonlinear Fourier coefficients are given by

$$a(\lambda) = \left(\cos(\Delta T_{\mathrm{rec}}) - \mathrm{j}\frac{\lambda}{\Delta}\sin(\Delta T_{\mathrm{rec}})\right)\mathrm{e}^{\mathrm{j}\lambda T_{\mathrm{rec}}}, \tag{4.56}$$

$$b(\lambda) = \frac{-A^*}{\Delta}\sin(\Delta T_{\mathrm{rec}})\mathrm{e}^{-\mathrm{j}\lambda T'_{\mathrm{rec}}}, \tag{4.57}$$

with $T'_{\mathrm{rec}} = t_2 + t_1$.

Using Eq. (4.35), the continuous spectrum of the rectangular pulse is

$$q_{\mathrm{c}}(\lambda) = A^*\mathrm{e}^{-2\mathrm{j}\lambda t_2}\left(\mathrm{j}\lambda - \Delta\cot(\Delta T_{\mathrm{rec}})\right)^{-1}. \tag{4.58}$$

The discrete eigenvalues can be found by setting Eq. (4.56) to zero and evaluating relation

$$\mathrm{j}\tan\left(T_{\mathrm{rec}}\sqrt{|A|^2 + \lambda_k^2}\right) = \sqrt{1 + \frac{|A|^2}{{\lambda_k}^2}}, \tag{4.59}$$

which has to be satisfied by all discrete eigenvalues of the rectangular pulse. Since there is some intricacy linked to the details of square roots of complex values, note that, in the definition of Δ as well as in Eq. (4.59) the square root of a complex number is the principal square root defined according to $\sqrt{z} = \sqrt{r}\cdot\mathrm{e}^{\mathrm{j}\frac{\varphi}{2}}$, where r is the amplitude and φ is the phase of complex number z. The discrete spectral amplitudes can be computed, using Eq. (4.36) and the λ-derivative of Eq. (4.56)

$$a_\lambda(\lambda_k) = \mathrm{j}T_{\mathrm{rec}}\left(\cos(\Delta T_{\mathrm{rec}})\Pi + \mathrm{si}(\Delta T_{\mathrm{rec}})\Pi\right)\mathrm{e}^{\mathrm{j}\lambda_k T_{\mathrm{rec}}}, \tag{4.60}$$

where $\Pi = 1 - {\lambda_k}^2/\Delta^2$ and $\mathrm{si}(x)$ is

$$\mathrm{si}(x) = \begin{cases} 1 & \text{for } x = 0 \\ \frac{sin(x)}{x} & \text{for } x \neq 0 \end{cases}. \tag{4.61}$$

In Fig. 4.3 the time domain pulse and its corresponding nonlinear spectrum is plotted for three different amplitudes. For the smallest amplitude $A = 1$ there are no

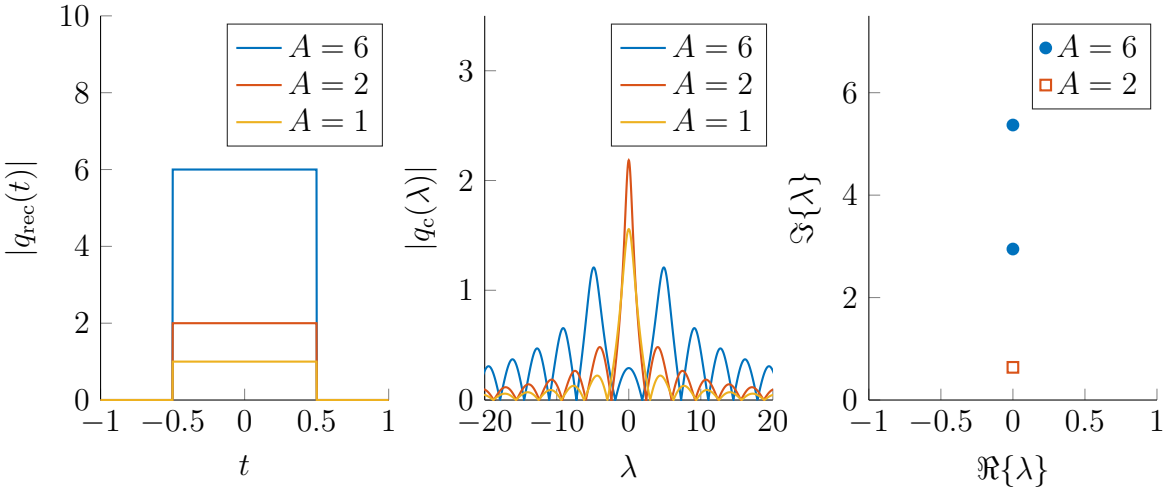


Figure 4.3.: Rectangular pulse and corresponding spectra for different values of A.

discrete eigenvalues. For higher values more and more discrete eigenvalues appear in the upper half complex plane.

4.3.2. The Fundamental Soliton

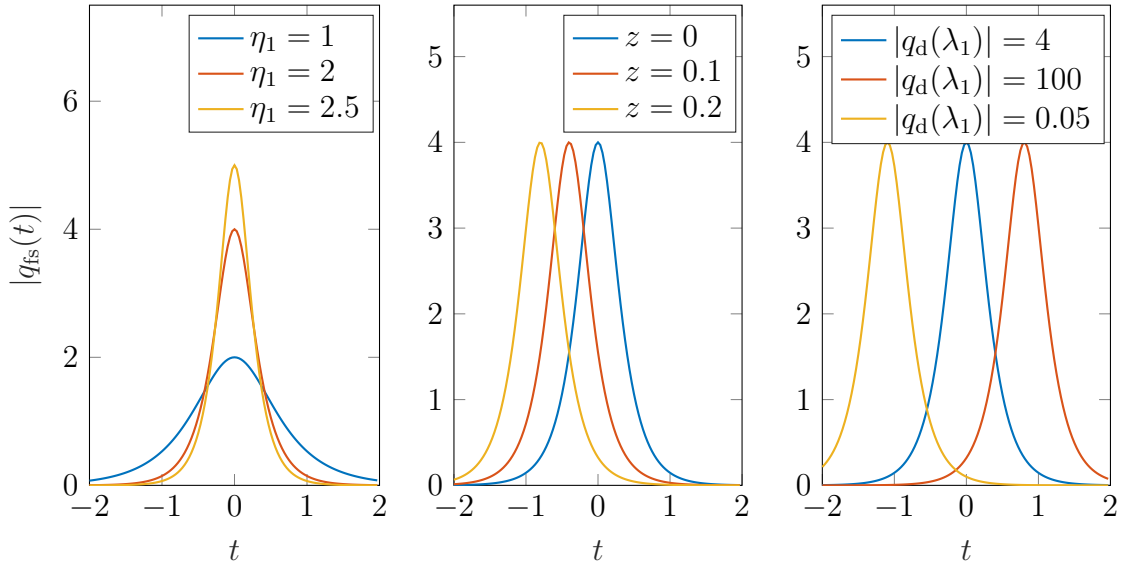
Another widely known pulse-shape, for which the nonlinear spectrum is known analytically, is the *fundamental soliton*. Its temporal description is [90], [103]

$$q_{\mathrm{fs}}(t,z) = -\mathrm{j}\mathrm{e}^{-\mathrm{j}\varphi_1}\mathrm{e}^{-\mathrm{j}4(\zeta_1^2-\eta_1^2)z}\mathrm{e}^{-\mathrm{j}2\zeta_1 t}2\eta_1 \cdot \operatorname{sech}\left(2\eta_1 t + 8\zeta_1\eta_1 z - \ln\left(\frac{|q_{\mathrm{d}}(\lambda_1)|}{2\eta_1}\right)\right), \quad (4.62)$$

where the only discrete eigenvalue is $\lambda_1 = \zeta_1 + \mathrm{j}\eta_1$. The corresponding discrete spectral amplitude is $q_{\mathrm{d}}(\lambda_1) = |q_{\mathrm{d}}(\lambda_1)|\mathrm{e}^{-\mathrm{j}\varphi_1}$. The above equation for the fundamental soliton can be retrieved, by evaluating the first iteration of the Darboux method, described in section 4.4.6. The center of this pulse is given by $t_0 = \left(\frac{1}{2\eta_1}\right)\ln\left(\frac{|q_{\mathrm{d}}(\lambda_1)|}{2\eta_1}\right) - 4\zeta_1 z$ and thus for a purely imaginary eigenvalue $\lambda_1 = \mathrm{j}\eta_1$, the pulse is centered at zero if $|q_{\mathrm{d}}(\lambda_1)| = 2\eta_1$.

In Fig. 4.4 the fundamental soliton is shown for several sets of parameters. Fig. 4.4a shows an ideally centered soliton with varying purely imaginary eigenvalues. Fig. 4.4b shows fundamental solitons with eigenvalues with a non-zero real part ($\zeta_1 = 1$)

for different values of z and Fig. 4.4c shows pulses with $\lambda_1 = \mathrm{j}2$ and varying values $|q_\mathrm{d}(\lambda_1)|$. It can be seen from Eq. (4.62) and Fig. 4.4, that fundamental solitons



(a) $|q_\mathrm{d}(\lambda_1)| = 2\eta_1$ (b) $\lambda_1 = 1{+}\mathrm{j}2$, $|q_\mathrm{d}(\lambda_1)| = 2\eta_1$ (c) $\lambda_1 = \mathrm{j}2$

Figure 4.4.: Fundamental Solitons

in general retain their shape during propagation. Fundamental solitons with purely imaginary eigenvalues ($\lambda_1 = \mathrm{j}\eta_1$) also have the additional property, that they do not drift in t during propagation (compare e.g. Fig. 4.4a and 4.4b). This property makes these pulses especially convenient to work with in simplified transmission scenarios and measurement algorithms as will be seen in e.g. section 4.10. Note, that the effects visible in Figs. 4.4b occur even though the denormalized temporal coefficient τ signifies retarded time, such that the usual co-propagating time-window representation is obtained in Eq. (3.1). This however only results in pulses that do not drift out of the allotted time window if the discrete eigenvalues used are purely imaginary. Thus, for e.g. the pulses in Fig. 4.4b, the soliton does drift from its initial position for $z \neq 0$ because of the non-zero real part. On the other hand, the pulses in Fig. 4.4c do not drift for $z \neq 0$. Instead, they retain their position in the time-slot, with their position being modulated using $|q_\mathrm{d}(\lambda_1)|$. This could e.g. be used to facilitate a pulse-position modulation (PPM) scheme with modulation in the nonlinear Fourier domain.

In general, it is not straightforward to develop analytic statements about the temporal and linear spectral width of pulses generated using the NFT. For the fundamental soliton however, it is feasible and the results are given in the following.

Temporal Width

The fundamental soliton has infinite temporal width, so an auxiliary width measurement is defined according to the percentage of energy p that is retained inside width T_p. The width can be determined using Eq. (4.49) and (4.62) in

$$E_{\mathrm{fs}}p = 4\eta_1 p = \int_{-T_p/2+t_0}^{T_p/2+t_0} |q_{\mathrm{fs}}(t,z)|^2 \mathrm{d}t, \tag{4.63}$$

where t_0 is the pulse-center and E_{fs} is the total pulse energy as given by Eq. (4.48). Solving this equation for temporal width T_p yields

$$T_p = \frac{\operatorname{arctanh}(p)}{\eta_1}. \tag{4.64}$$

A detailed derivation of this and the following properties is given in appendix A.3

Bandwidth

The bandwidth with respect to the standard Fourier transform can be obtained by similar means. Using Parsevals theorem [34] to state

$$E_{\mathrm{fs}}p = 4\eta_1 p = \int_{-B_p/2+f_0}^{B_p/2+f_0} |\mathcal{F}\{q_{\mathrm{fs}}(t,z)\}|^2 \mathrm{d}f, \tag{4.65}$$

where f_0 is the pulses linear center-frequency and solving for B_p, the relation

$$B_p = \frac{4\eta_1}{\pi^2}\operatorname{arctanh}(p), \tag{4.66}$$

can be obtained.

Time-Bandwidth Product

With Eqs. (4.64) and (4.66), the TBP can be computed by

$$\mathrm{TBP}(p) = T_p \cdot B_p = \frac{4}{\pi^2}\operatorname{arctanh}^2(p). \tag{4.67}$$

While temporal width and bandwidth are dependent on η_1, the TBP is independent of the imaginary part of the solitons discrete eigenvalue. Note, that for fundamental solitons a finite TBP can only exist for pulses truncated in time and linear frequency domain according to factor p. This differs from the definition used for linear modulation schemes using e.g. RRC pulses with matched filters, where the bandwidth is finite and the temporal width is defined such that the pulses, when sampled correctly do not interfere with each other under the assumption of an ideal channel. The definition for the linear modulation scheme can be used to answer questions regarding e.g. transceiver design more directly compared to the results obtained for fundamental solitons in Eqs. 4.64, 4.66 and 4.67. Nonetheless, the results are a useful starting point for the design of soliton transmission systems. A more detailed discussion of the practical use of the obtained parameters can be found in section 4.5, where a more practical metric for time support and bandwidth requirements for specific system setups is developed as well. Fig. 4.5 depicts the TBP of a fundamental soliton over the retained energy percentage p. A time-bandwidth product of TBP $= 1$ is reached for $p \approx 0.9172$.

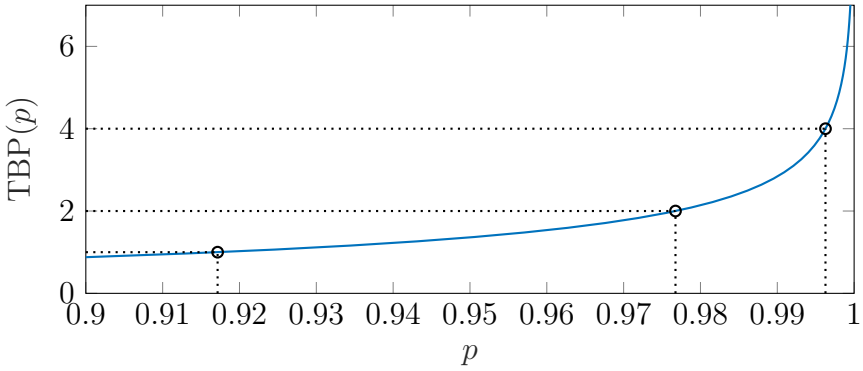


Figure 4.5.: Time-Bandwidth Product of Fundamental Soliton

For soliton based transmission systems, this is considered a rather low value, leading to considerable distortions of the corresponding nonlinear spectrum. Note, that the TBP shown in Fig. 4.5 is identical to the commonly used TBP for linear systems in the sense that it does not have to be denormalized, since the normalization coefficient (see Eq. (4.2)) is canceled out in Eq. (4.67). If it is assumed that multiple pulses are transmitted in neighboring time-slots, ISI is to be expected, due to the radiative components generated by truncation. Even a value of TBP $= 4$ results in $1 - p \approx 3 \cdot 10^{-3}$, which is still higher that values often assumed for p in literature. Since the SE of a system is directly proportional to the TBP, it can already be seen, that soliton based

transmission systems are unlikely to be able to compete with current state-of-the-art transmission systems with much lower TBPs.

4.3.3. The Satsuma-Yajima Pulse

The shape of a Satsuma-Yajima (SY) pulse in time-domain is given by [19], [104]

$$q_{\mathrm{sy}}(t) = A \cdot \operatorname{sech}(t). \tag{4.68}$$

While Eq. 4.68 at first glance might look like a special case of Eq. 4.62, it makes more sense to think of them as separate pulse-types, which only coincide if $A = 2\eta = 1$, $\varphi_1 = -\pi/2$ and $|q_{\mathrm{d}}(\lambda_1)| = 2\eta$. More generally, the SY pulse has discrete eigenvalues

$$\lambda_k = \mathrm{j}\left(A - k + \frac{1}{2}\right) \text{ for } \Im\{\lambda_k\} > 0. \tag{4.69}$$

and continuous nonlinear spectrum [27]

$$q_{\mathrm{c}}(\lambda) = \frac{\Gamma(-\mathrm{j}\lambda + \frac{1}{2} + A)\Gamma(-\mathrm{j}\lambda + \frac{1}{2} - A)}{\Gamma^2(-\mathrm{j}\lambda + \frac{1}{2})} \frac{\sin(\pi A)}{\cosh(\pi\lambda)}, \tag{4.70}$$

where $\Gamma(\cdot)$ is the complex gamma function [105]. Since solitonic pulses can not be linearly scaled, while fully maintaining their solitonic properties, for most choices of A the SY pulse also has a non-zero continuous nonlinear spectrum. According to Eq. (4.70) the spectrum of the SY pulse is purely discrete if $A \in \mathbb{N}$. In Fig. 4.6 two SY pulses and their corresponding continuous spectrum and discrete eigenvalues are depicted.

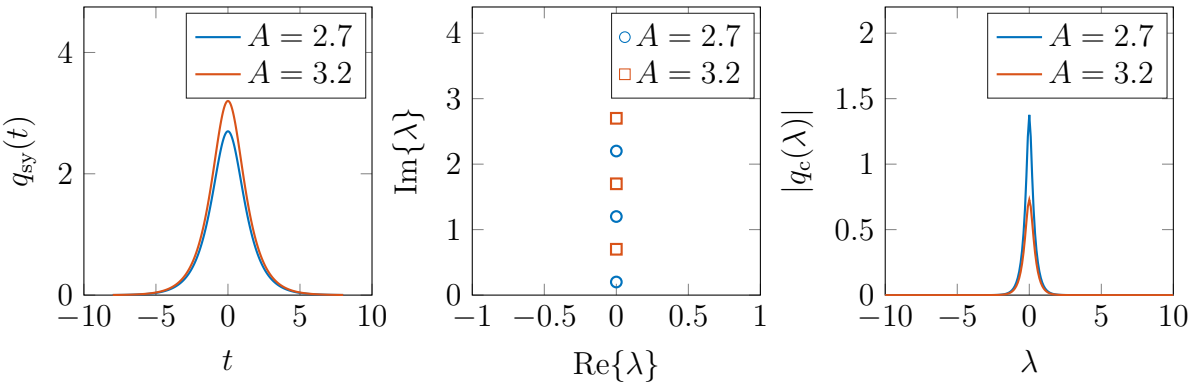


Figure 4.6.: SY pulses and their spectral components for varying amplitudes A.

4.4. Algorithms

For most pulse shapes, the NFT and its inverse can not computed analytically and thus numerical methods have to be developed, to enable NFDM transmission. A 'good' numerical algorithm for these transforms should exhibit low complexity, good numerical stability and high accuracy. The amount of existing algorithms is vast and has grown steadily for the recent years [27], [29], [106]. In the following, some examples from different classes of algorithms are presented. Some of the presented algorithms are only suitable for special cases or do only provide a subset of the nonlinear spectrum.

For the methods in the following sections, signal $q(t)$ is truncated to interval $[t_1, t_2]$ and discretized according to $q[n] = q(t_1 + nh)$, with index $n \in (0, N_\mathrm{s} - 1)$, h being the sample duration and N_s being the number of samples. The nonlinear Fourier coefficients will always be computed for the $t \to +\infty$ case, which for pulses on truncated temporal supports results in $t = t_2$.

4.4.1. Forward NFT for the Continuous Spectrum

In many cases, the continuous and discrete nonlinear spectrum are computed in independent steps. In this section, several algorithms to compute the continuous nonlinear spectrum are presented. They also, in many cases, form the basis for the computation of the discrete nonlinear spectrum, which will be treated in section 4.4.2. These two classes of algorithms, can be used jointly to obtain the complete NFT of an arbitrary pulse.

Forward Discretization Method

The forward discretization (FD) algorithm is based on the Euler scheme [107], which is a straightforward approach to numerical integration, for which the global error $e_\mathrm{FD} = \mathcal{O}(h)$ is proportional to the algorithms stepsize h. Using Eq. (4.17) under the assumption that, for each step, matrix $\boldsymbol{P}[n]$ and vector $\boldsymbol{v}[n]$ are constant, the iterative algorithm to compute the vector for time instances $t > t_1$ is [27]

$$\boldsymbol{v}[n+1] = \boldsymbol{A}_\mathbf{fd}[n]\boldsymbol{v}[n], \quad \boldsymbol{v}[0] = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathrm{e}^{-\mathrm{j}\lambda t_1}, \tag{4.71}$$

where $\boldsymbol{A}_{\mathbf{fd}}[n] = \boldsymbol{I}_{\mathbf{2}} + h\boldsymbol{P}[n]$, $\boldsymbol{I}_{\mathbf{2}}$ is the 2x2 identity matrix and

$$\boldsymbol{P}[n] = \begin{bmatrix} -\mathrm{j}\lambda & q[n] \\ -q^*[n] & \mathrm{j}\lambda \end{bmatrix}. \tag{4.72}$$

The initial value for the discretized vector $\boldsymbol{v}[0]$ is determined by the boundary condition at t_1, given by Eq. (4.22). The Fourier coefficients, and by extension the continuous spectrum $q_\mathrm{c}(\lambda) = b(\lambda)/a(\lambda)$, are obtained using Eq. (4.28). Thus, the nonlinear Fourier coefficients are approximated by

$$a(\lambda) \approx v_1[N_\mathrm{s} - 1]\mathrm{e}^{\mathrm{j}\lambda t_2}, \tag{4.73}$$

$$b(\lambda) \approx v_2[N_\mathrm{s} - 1]\mathrm{e}^{-\mathrm{j}\lambda t_2}, \tag{4.74}$$

where $v_1[n]$, $v_2[n]$ are the first and second entry of vector $\boldsymbol{v}[n]$ respectively.

Central Discretization

Improving on Eq. (4.71), the central discretization method (CDM) utilizes the central difference integration scheme, such that vector $\boldsymbol{v}[n]$ is computed by [27]

$$\boldsymbol{v}[n+1] = \boldsymbol{v}[n-1] + 2h\boldsymbol{P}[n]\boldsymbol{v}[n]. \tag{4.75}$$

Note that, in the first step the value for $\boldsymbol{v}[n]$ has to be computed by some other method e.g. FD. The continuous nonlinear spectrum can again be obtained from $\boldsymbol{v}[N_\mathrm{s} - 1]$, according to Eqs. (4.73) and (4.74).

Fourth-Order Runge-Kutta Method

The Runge-Kutta (RK) method uses a higher order integration scheme. The iterative update relation is [108]

$$\boldsymbol{v}[n+1] = \boldsymbol{v}[n] + \frac{h}{6}(\boldsymbol{r}_1 + 2\boldsymbol{r}_2 + 2\boldsymbol{r}_3 + \boldsymbol{r}_4), \tag{4.76}$$

with

$$\boldsymbol{r}_1 = \boldsymbol{P}[n]\boldsymbol{v}[n], \tag{4.77}$$

$$\boldsymbol{r}_2 = \boldsymbol{P}\left[n+\frac{1}{2}\right](\boldsymbol{v}[n] + \frac{h}{2}\boldsymbol{r}_1), \tag{4.78}$$

$$\boldsymbol{r}_3 = \boldsymbol{P}\left[n+\frac{1}{2}\right](\boldsymbol{v}[n] + \frac{h}{2}\boldsymbol{r}_2), \tag{4.79}$$

$$\boldsymbol{r}_4 = \boldsymbol{P}[n+1](\boldsymbol{v}[n] + h\boldsymbol{r}_3). \tag{4.80}$$

The matrix $\boldsymbol{P}[n+\frac{1}{2}]$ can be computed using the arithmetic middle between $q[n]$ and $q[n+1]$

$$\boldsymbol{P}\left[n+\frac{1}{2}\right] = \begin{bmatrix} -\mathrm{j}\lambda & \frac{1}{2}(q[n]+q[n+1]) \\ -(\frac{1}{2}(q[n]+q[n+1]))^* & \mathrm{j}\lambda \end{bmatrix}. \tag{4.81}$$

Note that there are other ways to obtain these intermediate values. However, since the algorithm already shows good accuracy using the above definition, other options are not investigated in this thesis.

Crank-Nicolson Method

In using the Crank-Nicolson (CN) method, the derivative in Eq. (4.17) is approximated by a finite difference scheme, which can be used with several different types of discretization. Here, CN refers to the approximation by FD, resulting in the relation [27]

$$\frac{\boldsymbol{v}[n+1]-\boldsymbol{v}[n]}{h} = \frac{1}{2}(\boldsymbol{P}[n+1]\boldsymbol{v}[n+1] + \boldsymbol{P}[n]\boldsymbol{v}[n]). \tag{4.82}$$

To facilitate numerical computation, the above equation can be made explicit

$$\boldsymbol{v}[n+1] = \left(\boldsymbol{I_2} - \frac{h}{2}\boldsymbol{P}[n+1]\right)^{-1}\left(\boldsymbol{I_2} + \frac{h}{2}\boldsymbol{P}[n]\right)\boldsymbol{v}[n] = \boldsymbol{A_{cn}}[n]\boldsymbol{v}[n]. \tag{4.83}$$

Ablowitz-Ladik Discretization

Using an integrable discrete version of the NLSE, introduced in [109], this scheme can be used to augment a variety of algorithms. Here, Ablowitz-Ladik (AL) will refer to the FD scheme that is augmented according to [109].

The iterative step is [27]

$$\boldsymbol{v}[n+1] = \begin{bmatrix} z & Q[n] \\ -Q^*[n] & z^{-1} \end{bmatrix} \boldsymbol{v}[n] = \boldsymbol{A}_{\mathbf{al}}[n]\boldsymbol{v}[n], \tag{4.84}$$

with $z = \mathrm{e}^{-\mathrm{j}\lambda h}$ and $Q[n] = q[n]h$. Since the continuous nonlinear spectrum is given by the ratio of the nonlinear Fourier coefficients, the above equation can be normalized to mitigate numerical errors. The normalized system is given by [110]

$$\boldsymbol{v}[n+1] = \frac{1}{\sqrt{1+|Q[n]|^2}} \begin{bmatrix} z & Q[n] \\ -Q^*[n] & z^{-1} \end{bmatrix} \boldsymbol{v}[n]. \tag{4.85}$$

Since it was found that, for the examples in this thesis, the difference between both methods is negligible, Ablowitz-Ladik method will refer to the first version, given in (4.84).

Trapezoidal Discretization

The last example for integration based methods is the trapezoidal discretization (TD) scheme. In a first step, Eq. (4.17) is modified, resulting in the new linear temporal propagation equation [106]

$$\boldsymbol{u}_t(t,\lambda) = \begin{bmatrix} 0 & q(t)\mathrm{e}^{\mathrm{j}2\lambda t} \\ -q^*(t)\mathrm{e}^{-\mathrm{j}2\lambda t} & 0 \end{bmatrix} \boldsymbol{u}(t,\lambda) = \boldsymbol{F}(t,\lambda)\boldsymbol{u}(t,\lambda), \tag{4.86}$$

for vectors $\boldsymbol{u}(t,\lambda)$ with boundary condition $\boldsymbol{u}(t_1,\lambda) = [1,0]^{\mathrm{T}}$ and $a(\lambda) = u_1(t_2,\lambda)$, $b(\lambda) = u_2(t_2,\lambda)$, where $u_1(t,\lambda)$ and $u_2(t,\lambda)$ are the entries of $\boldsymbol{u}(t,\lambda)$. This step can be facilitated by using $v_1(t,\lambda) = u_1(t,\lambda)\cdot\mathrm{e}^{-\mathrm{j}\lambda t}$ and $v_2(t,\lambda) = u_2(t,\lambda)\cdot\mathrm{e}^{\mathrm{j}\lambda t}$ in Eq. (4.17). This simplifies the computation of the matrix exponential in the following and also results in the nonlinear Fourier coefficients being directly available at the last step of the algorithm (see Eq. (4.89)). After discretization, the calculation of the nonlinear Fourier coefficients can be done by

$$\boldsymbol{w}[n+1] = \mathrm{e}^{\boldsymbol{F}[n]h}\boldsymbol{w}[n] = \boldsymbol{G}[n]\boldsymbol{w}[n], \tag{4.87}$$

where $\boldsymbol{w}[0] = \mathrm{e}^{\boldsymbol{F}[0]\frac{h}{2}}[1,0]^{\mathrm{T}} = \boldsymbol{G}^{\frac{1}{2}}[0][1,0]^{\mathrm{T}}$. Here, in an intermediate step, the initial solution for $\boldsymbol{u}(t_1,\lambda)$ is shifted by a halfstep. This is reversed in the step in Eq. (4.89)

to obtain the nonlinear Fourier coefficients. It can be seen, that this shift is what makes this integration scheme trapezoidal. The matrix exponential can be computed in closed form resulting in

$$\boldsymbol{G}[n] = \begin{bmatrix} \cos(|q[n]|h) & \sin(|q[n]|h)\mathrm{e}^{\mathrm{j}\varphi[n]+\mathrm{j}2\lambda t[n]} \\ -\sin(|q[n]|h)\mathrm{e}^{-\mathrm{j}\varphi[n]-\mathrm{j}2\lambda t[n]} & \cos(|q[n]|h) \end{bmatrix}, \tag{4.88}$$

with $\mathrm{e}^{\mathrm{j}\varphi[n]} = \frac{q[n]}{|q[n]|}$ and $t[n] = t + nh$. The nonlinear Fourier coefficients then can be obtained by using

$$\begin{bmatrix} a[N_\mathrm{s}-1] \\ b[N_\mathrm{s}-1] \end{bmatrix} = \boldsymbol{G}^{-\frac{1}{2}}[N_\mathrm{s}-1]\boldsymbol{w}[N_\mathrm{s}-1]. \tag{4.89}$$

The matrices $\boldsymbol{G}^{\pm\frac{1}{2}}[n]$ needed in the first and last step of the iteration are obtained by exchanging all stepsizes h in Eq. (4.88) by $\pm h/2$.

Layer-Peeling Method

The layer-peeling (LP) scheme is not based on numerically solving the ZS-system from Eq. (4.17). Instead, this scheme approximates the signal $q(t)$ by rectangular segments of width h, as exemplified in Fig. 4.7 and then uses the analytic expressions for the nonlinear Fourier coefficients from section 4.3.1 and the layer-peeling property given in section 4.2.2 to approximate the nonlinear Fourier coefficients of the pulse. The

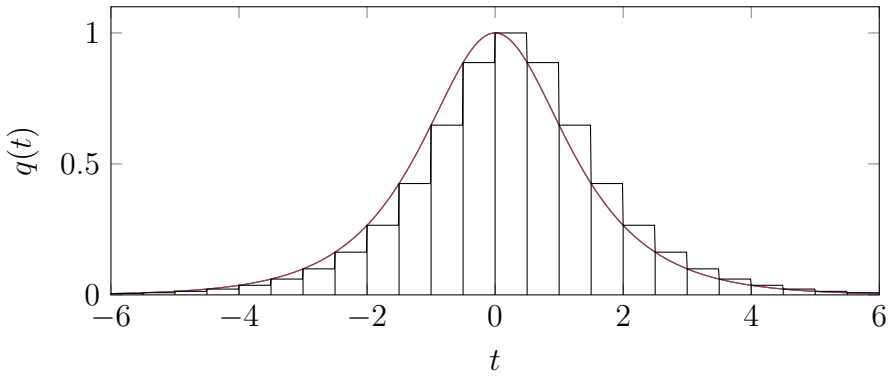


Figure 4.7.: Approximation of pulse by rectangular segments ($h = 0.5$)

iterative algorithm to obtain the nonlinear Fourier coefficients of the full pulse $q(t)$ is

[27]

$$a[n+1] = a[n]x[n] - b[n]\bar{y}[n], \tag{4.90}$$
$$b[n+1] = a[n]y[n] + b[n]\bar{x}[n],, \tag{4.91}$$

where the nonlinear Fourier coefficients $x[n]$, $y[n]$,$\bar{x}[n]$, $\bar{y}[n]$ of the rectangular pulse at timestep n are given by

$$x[n] = \left(\cos(Dh) - \mathrm{j}\frac{\lambda}{D}\sin(Dh)\right)\mathrm{e}^{\mathrm{j}\lambda h} \tag{4.92}$$
$$y[n] = \frac{-q^*[n]}{D}\sin(Dh)\mathrm{e}^{\mathrm{j}\lambda(t[n]+t[n-1])} \tag{4.93}$$
$$\bar{x}[n] = x^*[n](\lambda^*) \tag{4.94}$$
$$\bar{y}[n] = y^*[n](\lambda^*), \tag{4.95}$$

with $D = \sqrt{\lambda^2 + |q[n]|^2}$. The initial values for the nonlinear Fourier coefficients are $a[0] = 1$, $b[0] = 0$. This can be seen, if $q(t) = 0\ \forall\ \mathbb{R}$ is assumed in Eq. (4.17). The nonlinear Fourier coefficients of the whole pulse $q(t)$ are then approximated by $a(\lambda) \approx a[N_\mathrm{s} - 1]$, $b(\lambda) \approx b[N_\mathrm{s} - 1]$.

Accuracy Test for Rectangular Pulse

While the overall accuracy of a method is determined by the pulse-shape for which the transformation is to be executed, the closed form solutions from section 4.3 can be used to numerically evaluate the algorithms presented in this section for specific cases. The NMSEs (see Eq. (2.29)) between the analytic and numerical solutions for the above algorithms for a rectangular pulse with $A = 6$ and duration $T = 1$ are given in Fig. 4.8. It can be seen, that the continuous spectra $q_\mathrm{c}(\lambda)$ recovered by FD, CDM and AL exhibit a higher NMSE than the other methods. RK has superior accuracy but also is a higher order method and thus exhibits higher computational complexity. TD has similar complexity to the LP and FD methods, while being almost as accurate as RK. Note, that since there are no stochastic perturbations present in this measurement, the occurring error, while dependent on the used parameters and pulses, is not stochastic in nature. Thus, a single measurement of the NMSE for each data point in Fig 4.8 is sufficient. Similar measurements were done for three other pulses and can be found in appendix D.2. As previously noted these results do only apply to the specific cases

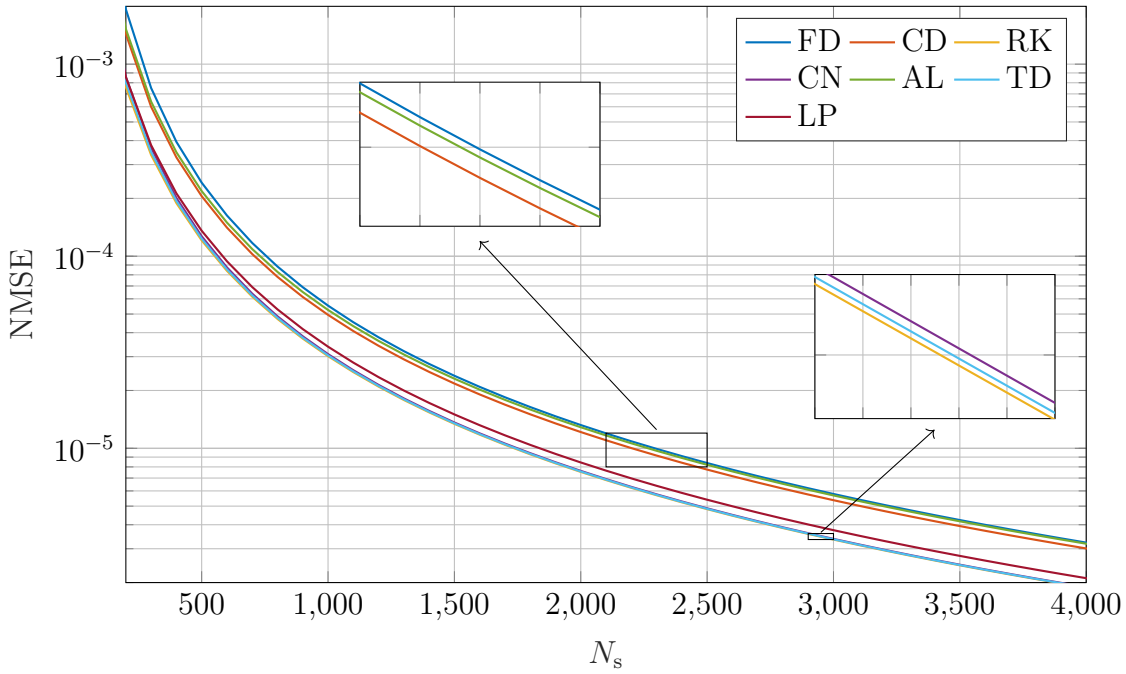


Figure 4.8.: Accuracy of continuous spectrum estimation for the rectangular pulse ($A = 6$, $T = 1$).

for which they were generated. Nonetheless, they give some insight into the numerical behavior of implementations of the presented algorithms.

4.4.2. Forward NFT for the Discrete Spectrum

Computing the discrete nonlinear spectrum, information about two different quantities has to be obtained, namely the discrete eigenvalues λ_k and the corresponding discrete spectral amplitudes $q_\mathrm{d}(\lambda_k)$. Alternatively, the nonlinear Fourier coefficient $b(\lambda_k)$ can be used directly, instead of $q_\mathrm{d}(\lambda_k)$. In general it is assumed, that the exact locations of the discrete eigenvalues are unknown and that there are either some reasonable guesses, or the area in which they are contained is known. Depending on the exact assumptions made, some of the presented algorithms are more suitable than others for specific cases. Computing the missing parameter $q_\mathrm{d}(\lambda_k)$ or $b(\lambda_k)$ for a determined λ_k value can be facilitated by extending the algorithms presented in 4.4.1. It has been shown, that in many cases it is advantageous to use the nonlinear Fourier coefficients $b(\lambda_k)$ instead of $q_\mathrm{d}(\lambda_k)$. However, all the algorithms presented in the following can also be used to obtain the discrete spectral amplitudes if needed.

Assuming, that some guesses $\lambda_k^{(0)}$ exist, which are close enough to the true values λ_k, these initial guesses can be used in a search-algorithm, to obtain estimates on the discrete eigenvalues of the respective pulse. In many transmission scenarios the assumption of having a set of guesses for the discrete eigenvalues is reasonable, since their locations might be part of the chosen modulation format. In such a case, the *Newton-Raphson* search can be used to find the zeros of $a(\lambda)$ in the upper-half complex plane. The search method updates the initial guess by [27, Sec. IV.A]

$$\lambda^{(i+1)} = \lambda^{(i)} - \alpha_{\mathrm{NR}} \frac{a(\lambda^{(i)})}{a_\lambda(\lambda^{(i)})}, \tag{4.96}$$

where $\lambda^{(i)}$ is the guess in the i-th iteration and step modifier α_{NR} is a scaling factor, that can be used to stabilize the search, by dampening the change in each step, if necessary. If not mentioned otherwise, the step modifier is set to $\alpha_{\mathrm{NR}} = 1$. The search is terminated if the change $\alpha_{\mathrm{NR}} a(\lambda^{(i)})/a_\lambda(\lambda^{(i)})$ becomes smaller than some threshold value δ_{NR}. It can be seen, that the values $a(\lambda^{(i)})$, $a_\lambda(\lambda^{(i)})$ are needed in each step. The first parameter can be obtained, using any method presented in section 4.4.1. The second parameter can be computed by extending the methods presented for the continuous spectrum. In the following, some of these extensions are reviewed shortly.

Forward Discretization Method

The FD method can be extended by calculating the λ-derivative of Eq. (4.71)

$$\boldsymbol{v}'[n+1] = \boldsymbol{A}'_{\mathbf{fd}}[n]\boldsymbol{v}[n] + \boldsymbol{A}_{\mathbf{fd}}[n]\boldsymbol{v}'[n], \quad \boldsymbol{A}'_{\mathbf{fd}}[n] = \begin{bmatrix} -\mathrm{j} & 0 \\ 0 & \mathrm{j} \end{bmatrix}, \quad \boldsymbol{v}'[0] = \begin{bmatrix} -\mathrm{j}t_1 \\ 0 \end{bmatrix} \mathrm{e}^{-\mathrm{j}\lambda t_1}, \tag{4.97}$$

and the derivative of Eq.(4.73)

$$a_\lambda(\lambda_k) \approx (v_1'[N_\mathrm{s}-1] + \mathrm{j}\lambda v_1[N_\mathrm{s}-1])\mathrm{e}^{\mathrm{j}\lambda t_2}. \tag{4.98}$$

After the search algorithm converged, an approximation of $b(\lambda_k)$ can be obtained, using Eq. (4.74) and an approximation of $q_\mathrm{d}(\lambda_k)$ can be obtained, by additionally using Eqs. (4.98) and (4.36).

Central Discretization

For the discrete spectrum CDM, Eq. (4.75) yields

$$\boldsymbol{v}'[n+1] = \boldsymbol{v}'[n-1] + 2h(\boldsymbol{P}'[n]\boldsymbol{v}[n] + \boldsymbol{P}[n]\boldsymbol{v}'[n]) \tag{4.99}$$

where $\boldsymbol{P}'[n]$ is identical to $\boldsymbol{A}'_{\mathbf{fd}}[n]$ in Eq. (4.97). The value for $\boldsymbol{v}'[n]$ in the first iteration, can again be computed by any other method e.g. by using Eq. (4.97).

Fourth-Order Runge-Kutta Method

For the RK algorithm the iteration step is given by

$$\boldsymbol{v}'[n+1] = \boldsymbol{v}'[n] + \frac{h}{6}(\boldsymbol{r}'_1 + 2\boldsymbol{r}'_2 + 2\boldsymbol{r}'_3 + \boldsymbol{r}'_4), \tag{4.100}$$

with

$$\boldsymbol{r}'_1 = \boldsymbol{P}'\boldsymbol{v}[n] + \boldsymbol{P}[n]\boldsymbol{v}'[n], \tag{4.101}$$

$$\boldsymbol{r}'_2 = \boldsymbol{P}'(\boldsymbol{v}[n] + \frac{h}{2}\boldsymbol{r}_1) + \boldsymbol{P}\left[n+\frac{1}{2}\right](\boldsymbol{v}'[n] + \frac{h}{2}\boldsymbol{r}'_1), \tag{4.102}$$

$$\boldsymbol{r}'_3 = \boldsymbol{P}'(\boldsymbol{v}[n] + \frac{h}{2}\boldsymbol{r}_2) + \boldsymbol{P}\left[n+\frac{1}{2}\right](\boldsymbol{v}'[n] + \frac{h}{2}\boldsymbol{r}'_2), \tag{4.103}$$

$$\boldsymbol{r}'_4 = \boldsymbol{P}'(\boldsymbol{v}[n] + h\boldsymbol{r}_3) + \boldsymbol{P}[n+1](\boldsymbol{v}'[n] + h\boldsymbol{r}'_3). \tag{4.104}$$

Note that $\boldsymbol{P}' = \boldsymbol{A}'_{\mathbf{fd}}[n]$ given in (4.97) is independent of n and thus the index is dropped in the equations above.

Crank-Nicolson Method

For the CN method the necessary additional parameter is given by [27]

$$\boldsymbol{v}'[n+1] = \boldsymbol{A}'_{\mathbf{cn}}[n]\boldsymbol{v}[n] + \boldsymbol{A}_{\mathbf{cn}}[n]\boldsymbol{v}'[n] \tag{4.105}$$

$$\boldsymbol{A}'_{\mathbf{cn}}[n] = \frac{h}{2}\left(\boldsymbol{I_2} - \frac{h}{2}\boldsymbol{P}[n+1]\right)^{-1}\begin{bmatrix} -\mathrm{j} & 0 \\ 0 & \mathrm{j} \end{bmatrix}(\boldsymbol{I_2} + \boldsymbol{A}_{\mathbf{cn}}[n]). \tag{4.106}$$

Ablowitz-Ladik Discretization

Using the AL discretization, the iteration step can be derived from Eq. (4.84) and is given by

$$\boldsymbol{v}'[n+1] = \boldsymbol{A}'_{\mathbf{al}}[n]\boldsymbol{v}[n] + \boldsymbol{A}_{\mathbf{al}}[n]\boldsymbol{v}'[n], \text{ with } \boldsymbol{A}'_{\mathbf{al}}[n] = h \begin{bmatrix} -\mathrm{j}z & 0 \\ 0 & \mathrm{j}z^{-1} \end{bmatrix}, \tag{4.107}$$

with $z = \mathrm{e}^{-\mathrm{j}\lambda h}$. The normalized version can be obtained by multiplying the right side of (4.107) by the normalization factor given in (4.85).

Trapezoidal Discretization

Similarly, for the derivatives $a_\lambda(\lambda)$, $b_\lambda(\lambda)$ the following set of equations can be derived [106]

$$\boldsymbol{w}'[n+1] = \boldsymbol{G}[n+1]\boldsymbol{w}'[n] + \boldsymbol{G}'[n+1]\boldsymbol{w}[n] \tag{4.108}$$

$$\boldsymbol{w}'[0] = (\boldsymbol{G}^{\frac{1}{2}})'[0] \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{4.109}$$

$$\begin{bmatrix} a'[N_\mathrm{s}-1] \\ b'[N_\mathrm{s}-1] \end{bmatrix} = \boldsymbol{G}^{-\frac{1}{2}}[N_\mathrm{s}-1]\boldsymbol{w}'[N_\mathrm{s}-1] + (\boldsymbol{G}^{-\frac{1}{2}})'[N_\mathrm{s}-1]\boldsymbol{w}[N_\mathrm{s}-1], \tag{4.110}$$

where

$$\boldsymbol{G}'[n] = \mathrm{j}2t[n]\sin(|q[n]|h) \begin{bmatrix} 0 & \mathrm{e}^{\mathrm{j}\varphi[n]+\mathrm{j}2\lambda t[n]} \\ \mathrm{e}^{-\mathrm{j}\varphi[n]-\mathrm{j}2\lambda t[n]} & 0 \end{bmatrix} \tag{4.111}$$

Again, the matrix $(\boldsymbol{G}^{\pm\frac{1}{2}})'[n]$ can be obtained by replacing h with $\pm\frac{h}{2}$ in the above equation. The derivatives of the nonlinear Fourier coefficients are approximated by $a_\lambda(\lambda) \approx a'[N_\mathrm{s}-1]$ and $b_\lambda(\lambda) \approx b'[N_\mathrm{s}-1]$.

Layer-Peeling Method

For the LP method the needed derivatives follow directly from (4.90-4.95) as

$$a'[n+1] = a'[n]x[n] + a[n]x'[n] - (b'[n]\bar{y}[n] + b[n]\bar{y}'[n]) \tag{4.112}$$

$$a'[n+1] = a'[n]y[n] + a[n]y'[n] + b'[n]\bar{x}[n] + b[n]\bar{x}'[n] \tag{4.113}$$

with

$$x'[n] = \mathrm{j}h\left(1 - \frac{\lambda^2}{D^2}\right)\left(\cos(Dh) - \frac{\sin(Dh)}{Dh}\right)\mathrm{e}^{\mathrm{j}\lambda h} \tag{4.114}$$

$$y'[n] = -q^*[n]\left\{\frac{\lambda h}{D^2}\cos(Dh) - \left(\frac{\lambda}{D^3} + \mathrm{j}\frac{t[n] + t[n-1]}{D}\right)\sin(Dh)\right\}\mathrm{e}^{-\mathrm{j}\lambda(t[n]+t[n-1])} \tag{4.115}$$

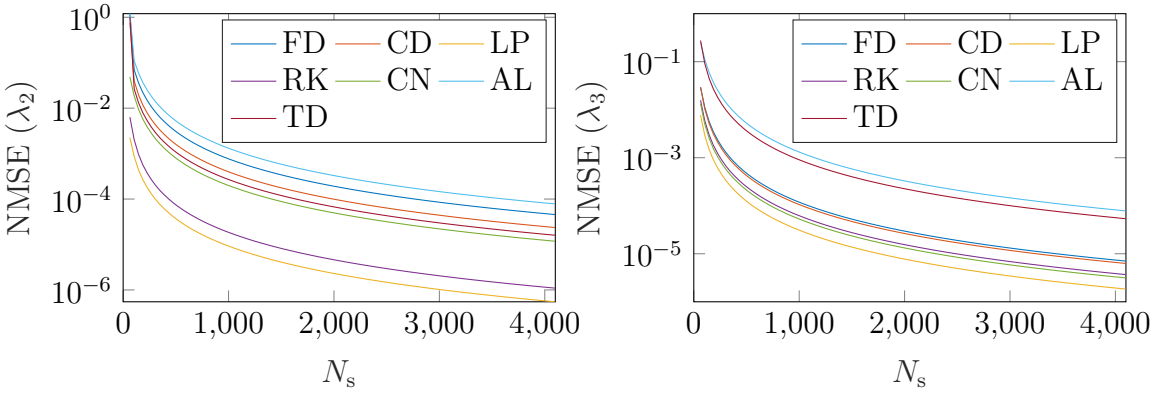
$$\bar{x'}[n] = (x')^*[n](\lambda^*) \tag{4.116}$$

$$\bar{y'}[n] = (y')^*[n](\lambda^*). \tag{4.117}$$

The initial values for the nonlinear Fourier coefficients derivatives $a'[0] = b'[0] = 0$ follow directly from the initial values given in section 4.4.1 and the derivative of the nonlinear Fourier coefficient is once more approximated by $a_\lambda(\lambda) \approx a'[N_\mathrm{s} - 1]$.

Accuracy Test for Satsuma-Yajima Pulse

A rough accuracy assessment for the methods described above can be done using the SY-pulse $q(t) = A\,\mathrm{sech}(t)$, described in section 4.3.3. For sufficiently high values $A \notin \mathbb{Z}^+$, the SY-pulse has both, continuous and discrete spectra. For integer values of A the continuous part vanishes and the spectrum is purely discrete. For $A = 2.7$, three discrete eigenvalues $\lambda_1 = \mathrm{j}0.2$, $\lambda_2 = \mathrm{j}1.2$, $\lambda_3 = \mathrm{j}2.2$ are present in the upper complex plane. Figure 4.9 shows the accuracy of the methods described above for the two largest eigenvalues, depending on the number of used samples.



(a) Estimation error for λ_2

(b) Estimation error for λ_3

Figure 4.9.: Accuracy of discrete spectrum NFT for SY-pulse ($A = 2.7$).

The used number of samples N_s range from 64 to 4096. Note, that similar to the measurements done for Fig 4.8, the occurring error is not stochastic in nature and just depends on the parameters and pulses used. Thus each data point in Figs. 4.9a and 4.9b was obtained by a single NMSE measurement. It can be seen, that in dependence on the magnitude of the imaginary part of the eigenvalue the resulting error for the used algorithms varies. Nonetheless, the LP algorithm produces the smallest error for all tested cases. As already stated in section 4.4.1, where the continuous spectrum NFT was studied, the error behavior is depending on the pulse-shape and other parameters, such as search-threshold δ_NR and thus, the values presented in Fig. 4.9 should only be taken as a rough estimate of the capabilities of the presented NFT algorithms.

Forward Backward Method

The methods described above, in many cases, generate a good estimate for the discrete eigenvalues λ_k, but tend to give dissatisfactory results for the associated nonlinear Fourier coefficients $b(\lambda_k)$ if N_s is small. Their performance can be improved, utilizing the forward-backward (FB) method proposed in [106]. In general, this method can be used with any NFT algorithm, for which the propagation matrices in t are unitary. In the following, the FB method will be discussed in light of the TD and LP schemes from section 4.4.1.

For the TD method, the spectral coefficients obtained from the iteration stated in Eq. (4.87) are given by [106]

$$\begin{bmatrix} a(\lambda) \\ b(\lambda) \end{bmatrix} \approx \underbrace{\boldsymbol{G}^{-\frac{1}{2}}[N_\mathrm{s}-1]\cdot\boldsymbol{G}[N_\mathrm{s}-1]\cdot\cdots\cdot\boldsymbol{G}[m+1]}_{\boldsymbol{R}}\cdot\underbrace{\boldsymbol{G}[m]\cdot\cdots\cdot\boldsymbol{G}[1]\boldsymbol{G}^{\frac{1}{2}}[0]}_{\boldsymbol{L}}\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{4.118}$$

$$= \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}\begin{bmatrix} L_{11} & L_{12} \\ L_{21} & L_{22} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{4.119}$$

It can be seen, that the nonlinear Fourier coefficients can now be approximated by

$$a[N_\mathrm{s}-1] = R_{11}L_{11} + R_{12}L_{21}, \tag{4.120}$$

$$b[N_\mathrm{s}-1] = R_{21}L_{11} + R_{22}L_{21}. \tag{4.121}$$

Additionally, it can be seen from Eq. (4.88) that $\boldsymbol{G}[n]$ and thus $\boldsymbol{R}$ and $\boldsymbol{L}$ are unitary

and thus the determinant

$$R_{11}R_{22} - R_{12}R_{21} = 1. \tag{4.122}$$

Using Eqs. (4.122) and (4.120), the estimate of nonlinear Fourier coefficient $b(\lambda_k)$, for some discrete eigenvalue λ_k can be given as

$$b[N_\mathrm{s}-1] = \underbrace{a[N_\mathrm{s}-1]\frac{R_{21}}{R_{11}}}_{0} + \frac{L_{21}}{R_{11}} = \frac{L_{21}}{R_{11}}. \tag{4.123}$$

Here the fact that $a[N_\mathrm{s}-1]$ should be zero if λ_k is a discrete eigenvalue was used to simplify the relation. It can be shown, that this increases the accuracy of the trapezoidal method, since matrix-entry R_{21} can become very large and thus amplify the potentially small numerical error in the estimation of $a(\lambda_k)$ [106].

Similarly, the FB method can be used for LP if the iteration from Eqs. (4.90) and (4.91) is rearranged to be

$$\begin{bmatrix} a[n+1] \\ b[n+1] \end{bmatrix} = \underbrace{\begin{bmatrix} x[n] & -\bar{y}[n] \\ y[n] & \bar{x}[n] \end{bmatrix}}_{\boldsymbol{M}[n]} \begin{bmatrix} a[n] \\ b[n] \end{bmatrix}. \tag{4.124}$$

The nonlinear Fourier coefficients are then defined by

$$\begin{bmatrix} a(\lambda) \\ b(\lambda) \end{bmatrix} \approx \underbrace{\boldsymbol{M}[N_\mathrm{s}-1]\cdot\dots\cdot\boldsymbol{M}[m+1]}_{\boldsymbol{R}} \cdot \underbrace{\boldsymbol{M}[m]\cdot\dots\cdot\boldsymbol{M}[0]}_{\boldsymbol{L}} \begin{bmatrix} a[0] \\ b[0] \end{bmatrix}, \tag{4.125}$$

where $a[0] = 1$ and $b[0] = 0$. It can also be shown, that the matrices $\boldsymbol{M}[n]$ are unitary and thus Eq. (4.123) can be directly applied to augment the LP-NFT using the FB method. Note that index m, at which the matrices are split up into $\boldsymbol{R}$ and $\boldsymbol{L}$, can be optimized [106]. Since this is beyond the scope of this thesis, the method will be used with $m = N/2$ over the course of the whole thesis.

In Fig. 4.10 the FB augmented TD method is compared with the standard TD scheme. It can be seen, that the NMSE for the b-values is improved by up to a factor of $\delta_\mathrm{NMSE} \approx 0.60$ for a low number of samples per pulse N_s, while for higher values of N_s the FB method does not significantly improve the accuracy in this particular case. For the evaluation, a fundamental soliton with $\lambda_1 = \mathrm{j}2$ and $b(\lambda_1) = -\mathrm{j}$ was generated,

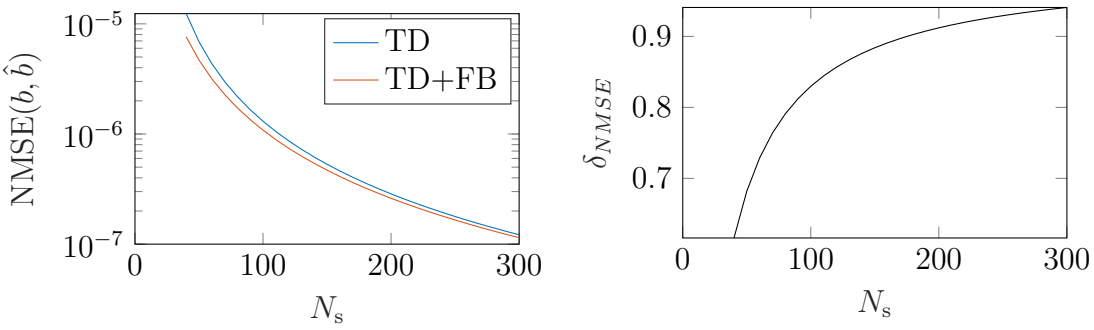


Figure 4.10.: Accuracy comparison for TD with and without FB method.

using the closed form formula. The pulse was evaluated using the respective NFT algorithm. Similar tests were done for the presented LP based FB method NFT, finding that the effect of the FB augmentation was almost negligible for all tested cases. While the factor R_{21}/R_{11} was comparable for both TD and LP based method, the deviation of $a[N_\mathrm{s} - 1]$ from zero was found to be smaller for all tested cases. It can thus be conjectured, that the step in Eq. (4.123) does change the accuracy less significantly for the LP based method. Note, that the evaluations in this section, were made without the addition of further perturbations such as ASE noise, so no statement on the change in robustness against perturbations can be made.

Central Difference Eigenproblem

In contrast to the previously described methods, the central difference eigenproblem (CDEP) method is only concerned with finding the discrete eigenvalues of a given pulse. To facilitate this, the discretized version of the signal is used to state a matrix eigenvalue problem, that can be solved, using any suitable numerical algorithm. The matrix eigenvalue problem is [27, Sec. IV.B]

$$\mathrm{j}\begin{bmatrix} \boldsymbol{D} & \mathrm{diag}(\boldsymbol{q}) \\ -\mathrm{diag}(\boldsymbol{q}^*) & -\boldsymbol{D} \end{bmatrix} \boldsymbol{v} = \mathbf{L}_{\mathbf{CD}}\boldsymbol{v} = \lambda \boldsymbol{v}, \tag{4.126}$$

with $\boldsymbol{q}$ being the vector containing the samples of $q(t)$, diag$(\cdot)$ being the diagonal matrix operator and

$$\boldsymbol{D} = \frac{1}{2h}\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & -1 \\ -1 & 0 & 1 & \cdots & 0 & 0 \\ & & & \cdots & & \\ 0 & 0 & -1 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & -1 & 0 \end{bmatrix}, \tag{4.127}$$

being the *central finite difference matrix* [27]. The discrete eigenvalues are then obtained by using some matrix eigenvalue method on $2N_\mathrm{s}$x$2N_\mathrm{s}$ matrix $\mathbf{L_{CD}}$. While the exact performance of the method depends on the algorithm used to calculate the eigenvalues of the matrix, it can be seen that in any case the complexity increases with the number of samples per pulse. Additionally, some algorithms will find all $2N_\mathrm{s}$ eigenvalues of the matrix. The additional eigenvalues found close to the real axis are often called *spurious eigenvalues* and have to be filtered by some additional step.

In Fig. 4.11 the unfiltered eigenvalues of the above described matrix are depicted for a Satsuma-Yajima pulse with $A = 2$. It can be seen, that many spurious eigenvalues

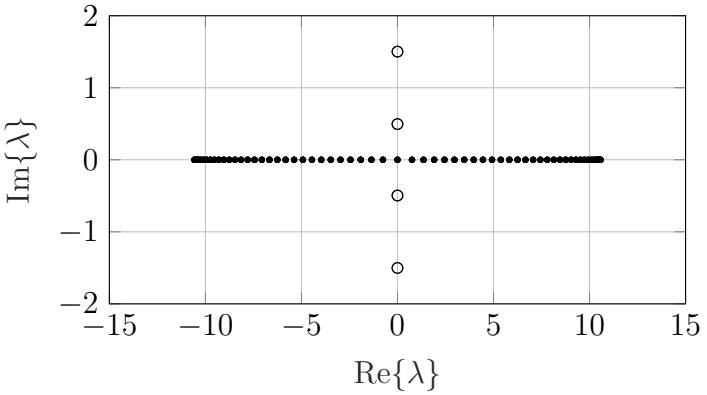


Figure 4.11.: Matrix eigenvalues for the Satsuma-Yajima pulse ($A = 2$).

are generated along the real axis and, due to the symmetry of discrete eigenvalues mentioned in section 4.1.5, the expected discrete eigenvalues at $\lambda_1 = \mathrm{j}0.5$ and $\lambda_2 = \mathrm{j}1.5$ are accompanied by their complex conjugates. Apart from the described method, there are other matrix base methods to compute the discrete eigenvalues. Some of them can be found in [27, Sec. IV.B].

Phase Jump Tracking

As mentioned at the beginning of section 4.4.2, the presented algorithms have varying applicability, depending on the assumptions made on the system. If e.g. it is assumed, that, due to the chosen modulation format, the ideal positions of the discrete eigenvalues are known, the Newton-Raphson search based algorithms are often quite efficient. On the other hand, if there is no information on the amount of discrete eigenvalues and their approximate location, the CDEP method can be used, to reliably find the discrete eigenvalues in exchange for a rather high complexity.

The phase jump tracking (PJT) method presented below is best applied if the location and number of discrete eigenvalues is unknown, but the region in which all λ_k lie is known. PJT is based on the argument principle [111]

$$K - P = \frac{1}{2\pi \mathrm{j}} \int_{\partial G} \frac{\mathrm{df}}{\mathrm{f}} = \frac{1}{2\pi} \Delta_{\partial G} \arg(\mathrm{f}). \tag{4.128}$$

Here K and P are the number of zeros and poles in domain G, f is some function, ∂G is some closed non-self-intersecting contour around domain G and $\Delta_{\partial G}\arg(\cdot)$ is the change in the argument of function f along contour ∂G. Using $a(\lambda)$ as the function in (4.128) results in

$$\Delta_{\partial G} \arg(a(\lambda)) = 2\pi K, \tag{4.129}$$

since $a(\lambda)$ is analytic in $\mathbb{C}^+$, as long as the contour ∂G lies in $\mathbb{C}^+$ as well. Further, there have to be K jumps in the argument of $a(\lambda)$ along ∂G. These jumps extend into the domain along trajectories ending at the discrete eigenvalues [111]. Finding the jumps along ∂G and following the trajectories is the core concept of PJT.

The contour ∂G has an upper border, that can be set by $U_{\partial G} = C_\mathrm{u} \cdot 0.25(E_\mathrm{t} - E_\mathrm{c})$, where E_t and E_c are the energy in time domain and nonlinear continuous spectrum respectively. The parameter $C_\mathrm{u} > 1$ is some constant, that can be used to extend domain G upwards to account for the movement of discrete eigenvalues away from the real axis, due to e.g. the addition of noise in time domain during propagation. If $U_{\partial G}$ is chosen too small, some discrete eigenvalues might lie outside of G and thus can not be found. On the other hand, if $U_{\partial G}$ is too high, either accuracy or computation time will be negatively affected, due to the increased length of $U_{\partial G}$. In the following, the parameter will be fixed to $C_\mathrm{u} = 1.1$. The left and right boundaries $L_{\partial G}$ and $R_{\partial G}$ can

be set according to [111]

$$\frac{|\hat{Q}(\omega_{L/R})|^2}{\hat{Q}_{\max}} = C_{\mathrm{LR}} \tag{4.130}$$

with $\hat{Q}(\omega)$ being the linear Fourier transform of the signal and $\hat{Q}_{\max} = \max_{\omega} |\hat{Q}(\omega)|^2$. The boundary values are then given by the minimum frequency $L_{\partial G} = \omega_L$ and maximum frequency $R_{\partial G} = \omega_R$, for which the above equation is satisfied. Note that presumably this is based on the observation that linear and nonlinear bandwidth are linked, assuming that this approximation still works if C_{LR} is small enough. Following [111], the constant is set to $C_{\mathrm{L/R}} = 10^{-4}$ in the following. The contour ∂G is then given by $G = [L_{\partial G}, R_{\partial G}] \times [\epsilon, U_{\partial G}]$, where the lower boundary ϵ is chosen to be a small positive value to prevent the detection of spurious discrete eigenvalues close to the real line.

The contour is then discretized by I equidistant points z_i with $i \in [1, I]$ on the contour. If a phase-jump in nonlinear Fourier coefficient $a(z_i)$ between two adjacent points occurs, these positions are stored for evaluation in the next step. The pseudo-code for this step is given in Algorithm 1.

Algorithm 1: Phase jump detection on ∂G (adapted from [111])

Input: $\{z_i\} \in \partial G$ - ordered points on the boundary
Output: $\{g_j\}$ - list of phase jump positions
for $j = 2, \cdots, I$ **do**
  if $\widetilde{\arg}(a(z_i)) \cdot \widetilde{\arg}(a(z_{i-1})) < 0$ and $|\widetilde{\arg}(a(z_i)) - \widetilde{\arg}(a(z_{i-1}))| > (c_{\mathrm{pj}} \cdot \pi)$ **then**
    add $\frac{z_i + z_{i-1}}{2}$ to $\{g_j\}$
  end
end

Here, function $\widetilde{\arg}(\cdot) = \arg(\cdot) - \pi$ is the well known $\arg(\cdot)$ function shifted by π, such that the new range of the function is $\widetilde{\arg}(\cdot) \in [-\pi, \pi)$. Note, that the value $c_{\mathrm{pj}} = 1.3$ and is a design parameter and was set heuristically according to [111]. The algorithm adds phase jump positions to set $\{g_j\}$ if two adjacent points z_i and z_{i-1} generate different signs and differ by an absolute value of more than $c_{\mathrm{pj}} \cdot \pi$ if evaluated by $\widetilde{\arg}(\cdot)$. Changes of sign occur at phase-jumps and around $\widetilde{\arg}(a(z_i)) = 0$. Checking for a certain minimum phase difference, ensures that not every change of sign is registered as a phase-jump. For small enough steps, this value does not play a significant role. For a very coarse set of $\{z_i\}$ the chosen value c_{pj} can result in phase-jumps not being

found reliably. In the conducted tests, using the heuristic value $c_{\mathrm{pj}} = 1.3$ was never an issue for a reasonably high number of values in $\{z_i\}$. The positions of the discrete eigenvalues can then be found by following the phase jump from its point on the boundary ∂G to some point at which it vanishes. In this point $a(\lambda) = 0$ and thus a discrete eigenvalue is found. There are a lot of subtleties to determining the step size for this process, optimizing for a good ratio between accuracy and complexity, which are detailed in [111]. The phase jump trajectories can be traversed by choosing a suitable algorithm. In the following, the *marching-squares algorithm* [112] is used and the pseudo-code for an implementation is given in appendix B.3. Depending on the chosen resolution for the marching-squares algorithm, the results have to be further refined by e.g. an additional search step, which is assumed to converge quickly and thus cause little additional overhead. A depiction of the points traversed by the algorithm for a SY-pulse is given in Fig. 4.12.

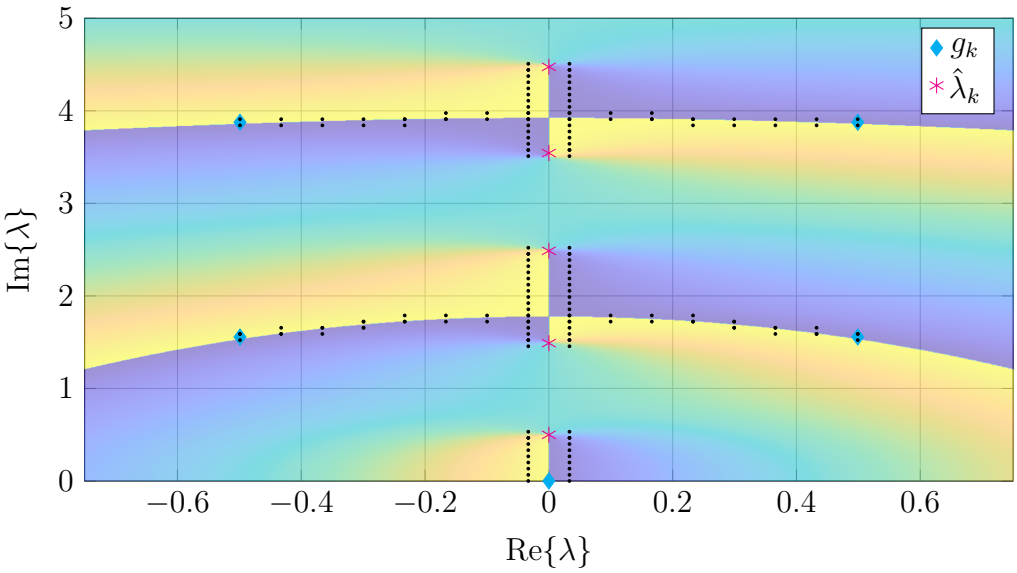


Figure 4.12.: Execution of the PJT algorithm for a SY-pulse ($A = 7$). Not all eigenvalues shown. Plane is colored according to the phase $\widetilde{\arg}(a(\lambda))$ (Yellow $\rightarrow -\pi$, Blue $\rightarrow +\pi$).

The algorithm first determines the points g_k on ∂G using Algorithm 1 and then uses a marching-squares algorithm to follow the phase-jump of $a(\lambda)$ to the estimates $\hat{\lambda}_k$. At the crossing points of the phase-jump trajectories on the imaginary axis, the marching squares algorithm has to follow one of the lines. In general, the rule to choose one of the branching paths can be chosen arbitrarily. Here it is assumed, that the algorithm

will always follow the next path in clockwise direction. If a further refinement step is used, the accuracy of this method is primarily determined by the refinement method. An in-depth analysis of the capabilities of this method can be found in [111].

4.4.3. The Eigenvalue Removal NFT

Assuming the NFT is to be used on some signal that contains continuous nonlinear spectral components, as well as a non-empty discrete spectrum, the algorithms previously described in this thesis can be used independent of each other to recover the respective parts of the nonlinear spectrum. The eigenvalue removal (ER) method adds one more step to iteratively remove computed discrete spectrum components of $q(t)$, before computing the remaining quantities. This method was first described in [113] and subsequently evaluated for full nonlinear spectrum pulses in [114], [115].

For the ER-method, the discrete spectrum method has to be chosen such that a single discrete eigenvalue can be estimated effectively. In the following, a search-based method with a single initial guess will be used. The ER-augmented search iteration then includes the standard search method, now followed by a subsequent removal of the estimated eigenvalue $\hat{\lambda}_k$ from $q^{(i)}(t)$. Additionally the pulse is then truncated if possible. A diagram for the 'search, remove, truncate' step is presented in Fig. 4.13. The received pulse is $q^{(1)}(t)$ and in each iteration pulse $q^{(i)}(t)$ is first used in a search

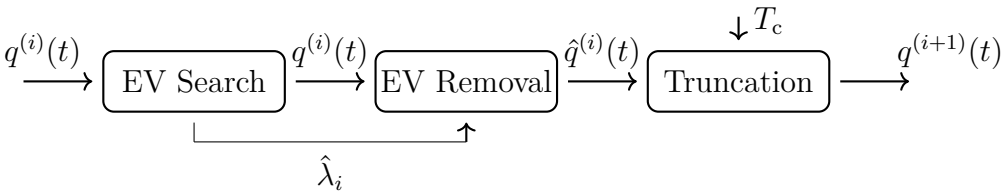


Figure 4.13.: Block diagram for one 'search, remove, truncate' step.

step to obtain estimate $\hat{\lambda}_i$, which is subsequently used to remove the corresponding time domain components from the used pulse. The resulting pulse $\hat{q}^{(i)}(t)$ no longer has $\hat{\lambda}_i$ as a discrete eigenvalue. In the final step of the iteration, this new pulse is truncated to temporal length $T_p^{(i)}$ according to some energy percentage constraint p. This truncated pulse $q^{(i+1)}(t)$ is then used in the search-method of the next iteration $i+1$. The estimated eigenvalue $\hat{\lambda}_i$ is removed from the temporal pulse, by using the

relation [113]

$$\hat{q}^{(i)}(t) = q^{(i)}(t) + \frac{2\mathrm{j}(\hat{\lambda}_i^* - \hat{\lambda}_i)v_2^*(t,\hat{\lambda}_i)v_1(t,\hat{\lambda}_i)}{|v_1(t,\hat{\lambda}_i)|^2 + |v_2(t,\hat{\lambda}_i)|^2}. \tag{4.131}$$

Multi-solitons are not constructed by linear superposition of fundamental solitons. However, the observations made in [116], [117] still justify the removal of discrete eigenvalues in order of increasing imaginary parts. Note that for each removal step the vectors $v_1(t,\hat{\lambda}_i)$ and $v_2(t,\hat{\lambda}_i)$ have to be computed. Also the removal step might fail if the estimate $\hat{\lambda}_i$ deviates strongly from the real value λ_i. In this case, a new discrete eigenvalue $\hat{\lambda}_i$ is added to $\hat{q}^{(i)}(t)$. This error can be monitored, by determining the pulse energy before and after the removal step, which should be reduced by $\approx 4\Im(\hat{\lambda}_i)$ if the removal was successful.

If the signal component corresponding to the continuous nonlinear spectrum is temporally more confined than the discrete spectrum part, the temporal support of the modified pulse $\hat{q}^{(i)}(t)$ can often be reduced significantly, while still fulfilling some in-window energy criterion (e.g. 99.99% energy inside truncated support $[t_1^{(i)}, t_2^{(i)}]$). This reduces the amount of processed samples in future steps and thus reduces the algorithms overall complexity. Assuming that all the removal steps are successful, the remaining time-domain signal $\hat{q}_{seed}(t)$ only contains continuous nonlinear spectral components and can be transformed by any of the continuous spectrum NFTs given in section 4.4.1. Note that parameter T_c, introduced in Fig. 4.13, is a lower bound on truncation. If the energy in the continuous spectrum is much smaller compared to the energy in the discrete spectrum, the truncation according to some in-window energy criterion might affect the continuous spectrum more strongly. To mitigate the potential degradation of the continuous spectrum, parameter T_c can be introduced as a minimum value for the signals support. This of course also restricts the potential reduction of complexity and thus T_c has to be set carefully. The choice of T_c has to be made heuristically. A good reference for the choice of T_c is the average temporal width of the pulse $\hat{q}_\mathrm{seed}(t)$, containing only the continuous nonlinear spectrum. This value can be measured at the transmitter for a fixed pulse alphabet and has to be known at the receiver during detection. A "blind" approach could be to use the eigenvalue removal algorithm without truncation for a training set of a certain size. Since there are methods to check for the successful removal of discrete spectrum parts, the remaining continuous spectrum only pulse can be obtained with a fairly high certainty. Since the pulses are not truncated yet, a value T_c can be computed from the average temporal support of pulses $\hat{q}_\mathrm{seed}(t)$ according to some factor p. For

a successful removal of an eigenvalue, the continuous spectrum is altered according to equation (4.163), which has to be taken into account after the transformation. The

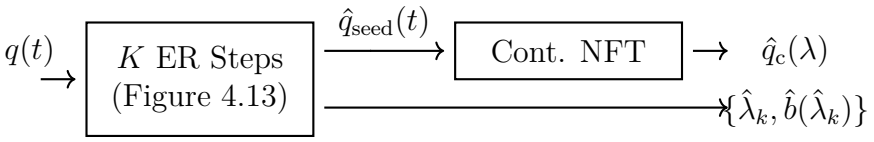


Figure 4.14.: Block diagram of the ER NFT.

full eigenvalue removal NFT algorithm is also depicted in Fig. 4.14, where $\hat{q}_{\text{c}}(\lambda)$ is the continuous spectrum estimate and $\hat{b}(\hat{\lambda}_k)$ is the estimate of the discrete spectrum nonlinear Fourier coefficient. Once more, an estimate of $q_{\text{d}}(\lambda_k)$ could have been used as well.

An example of this deconstruction, for a pulse with two discrete eigenvalues and a non-zero continuous spectrum, is given in Fig. 4.15. Note that the truncation step is skipped for the sake of clarity. The nonlinear spectrum of the initial pulse in Fig. 4.15a

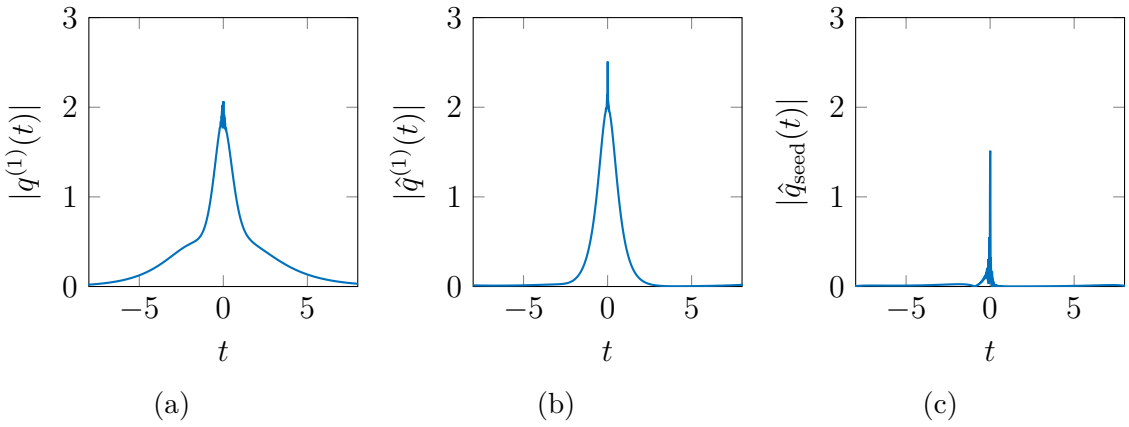


Figure 4.15.: Pulses in eigenvalue removal NFT iterations.

consists of the continuous spectrum $q_{\text{c}}(\lambda)$, the discrete eigenvalues $\{\lambda_1 = \text{j}0.25, \lambda_2 = \text{j}\}$ and corresponding $\{b(\lambda_1) = 1, b(\lambda_2) = \text{j}\}$. In Fig. 4.15b, the discrete eigenvalue with the smallest imaginary part λ_1 has been removed. The continuous spectrum is still present in the pulse, but is now distorted according to the relation in Eq. (4.163). The discrete spectrum includes only the remaining $\lambda_2 = \text{j}$. In Fig. 4.15c, λ_2 has been removed as well. The resulting continuous spectrum pulse does not have any discrete eigenvalues.

Complexity

The complexity of the discussed NFT algorithms depends on several factors. In addition to the number of samples N used to discretize the time domain pulse $q(t)$, the specific algorithm for each calculation of the nonlinear Fourier coefficients $a(\lambda)$, $b(\lambda)$ factors into the overall complexity. To obtain the nonlinear Fourier coefficients for a nonlinear frequency λ, $\mathcal{O}(N)$ multiplications and additions have to be executed [27]. Thus, the complexity of the continuous spectrum NFT is given by $\mathcal{O}(N^2)$, assuming the continuous spectrum is calculated on a N-point grid as well.

The complexity of one search iteration of the discrete search-based method, is $\mathcal{O}(N)$. The overall complexity also depends on the number of discrete eigenvalues K and the average number of search iterations for each discrete eigenvalue, which depends on the transmission channel, the transmitted pulse shape, the step modifier α_{NR} and the initial guesses. For the ER algorithm, the complexity for one search iteration can be expressed by $\mathcal{O}(C_{\mathrm{t}}(\lambda_k)N)$, where $C_{\mathrm{t}}(\lambda_k) \leq 1$ with $k \in [1, K]$ is used to include the reduction of signal samples after successful truncation. The overall complexity of the discrete spectrum part depends on the number of discrete eigenvalues K and the number of search iterations for each discrete eigenvalue. For a complexity comparison of the search-based and the ER-NFT, the number of search steps are assumed to be approximately equal, since the same received pulse $q(t)$ is used. Truncation is limited by T_{c} and thus $C_{\mathrm{t}}(\lambda_k) \in [T_{\mathrm{c}}/T_{\mathrm{S}}, 1]$. It can be seen, that if $C_{\mathrm{t}}(\lambda_k) = 1 \,\forall\, k \in [1, K]$ the complexity of all search steps is $\mathcal{O}(N)$ and thus the complexity for the discrete spectrum NFT is identical to the search-based method discrete spectrum NFTs complexity. The complexity of the continuous spectrum NFT is reduced if the truncated pulse $\hat{q}_{\mathrm{seed}}(t)$ is used as shown in Fig. 4.14. Its complexity is then given by $\mathcal{O}(C_{\mathrm{t}}(\lambda_K)N^2)$. Additionally, we would like to note that, since the continuous NFT can be implemented using any algorithm, fast NFT approaches are applicable, potentially reducing the complexity to $\mathcal{O}(C_{\mathrm{t}}(\lambda_K)N \log^2(C_{\mathrm{t}}(\lambda_K)N))$ [31]. Since the values for $C_{\mathrm{t}}(\lambda_k)$ depend on the specific shape of the received pulse and the width of the pulse in time domain, which cannot be linked to the nonlinear Fourier spectrum analytically for $K > 1$, the complexity of the search- and ER-based algorithms can only be compared in more detail by computing the values $C_{\mathrm{t}}(\lambda_k)$ numerically for some specific cases.

Numerical Evaluation

Tests of the INFT/NFT configuration in a back-to-back setup were conducted for pulses with $N = 2^{10}$ samples. The continuous nonlinear spectrum was modulated, using one nonlinear RRC shaped channel centered around $\lambda = 0$ with nonlinear spectral width $W_\lambda = 14.2857$ and amplitude A. The continuous nonlinear frequency support was set according to [23, eq. 22 f]. This channel was used to transmit phase-shift keying (PSK) symbols a_i with i.i.d. uniform random phase $\varphi_i \in [0, 2\pi)$. Note, that it is not yet well established how the nonlinear and linear spectral width of a pulse are related. While there are some results for limit cases [43], they are only strictly valid if the power of a pulse is going to zero. Thus, the width of the receiver filter has to be set heuristically. The discrete spectrum was either modulated using $K = 2$ or $K = 3$ discrete eigenvalues $\{\lambda_k = \mathrm{j}kC_{\mathrm{EV}}\}$. The corresponding $b(\lambda_k)$-values were again modulated with an PSK constellation with i.i.d. uniform random phases. Time domain transmission pulses were generated, using the INFT algorithm presented in section 4.4.7. The nonlinear spectrum of the generated pulses was obtained by either separate continuous or discrete NFTs or by the joint ER-based NFT described above. For comparison of the two NFT algorithms, the NMSEs between the modulated values $\{a_i$, $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_K]$, $\mathbf{b} = [b(\lambda_1), \cdots, b(\lambda_K)]\}$ and recovered values $\{\hat{a}_i$, $\hat{\boldsymbol{\lambda}} = [\hat{\lambda}_1, \cdots, \hat{\lambda}_K]$, $\hat{\mathbf{b}} = [\tilde{b}(\hat{\lambda}_1), \cdots, \tilde{b}(\hat{\lambda}_K)]\}$ were computed. Here the i-index signifies that a_i belongs to the i-th transmitted pulse and $\boldsymbol{a}$, $\hat{\boldsymbol{a}}$ are the vectors containing the symbols a_i and $\hat{a}_i$ respectively. The NMSEs are denoted as $e_{\mathrm{x}}(\boldsymbol{a}, \hat{\boldsymbol{a}})$ with $\mathrm{x} \in \{\mathrm{NR}, \mathrm{ER}\}$ for the NMSE calculated for continuous spectrum symbols, with the subscript NR indicating that the search-based NFT was used and similarly, ER indicating that the ER-NFT was used. The NMSE for the discrete eigenvalues is denoted as $e_{\mathrm{x}}(\boldsymbol{\lambda}, \hat{\boldsymbol{\lambda}})$ and the NMSE for the $b(\lambda_k)$-values is denoted as $e_{\mathrm{x}}(\boldsymbol{b}, \hat{\boldsymbol{b}})$. The minimum and maximum values obtained in the simulations are given in Table 4.1.

	Cont. Spectrum		Disc. Eigenvalues		b-Values	
	min	max	min	max	min	max
e_{NR} $(K = 2)$	$5.3 \cdot 10^{-5}$	$1 \cdot 10^{-2}$	$5 \cdot 10^{-7}$	$4 \cdot 10^{-4}$	$6.2 \cdot 10^{-7}$	$7.9 \cdot 10^{-5}$
e_{ER} $(K = 2)$	$5.3 \cdot 10^{-5}$	$1 \cdot 10^{-2}$	$3.8 \cdot 10^{-7}$	$2 \cdot 10^{-4}$	$5.2 \cdot 10^{-7}$	$7.6 \cdot 10^{-5}$
e_{NR} $(K = 3)$	$5.3 \cdot 10^{-5}$	$1 \cdot 10^{-2}$	$3.5 \cdot 10^{-6}$	$2.9 \cdot 10^{-3}$	$5.2 \cdot 10^{-6}$	$1 \cdot 10^{-3}$
e_{ER} $(K = 3)$	$5.3 \cdot 10^{-5}$	$1 \cdot 10^{-2}$	$3 \cdot 10^{-6}$	$2.7 \cdot 10^{-3}$	$4.2 \cdot 10^{-6}$	$1 \cdot 10^{-3}$

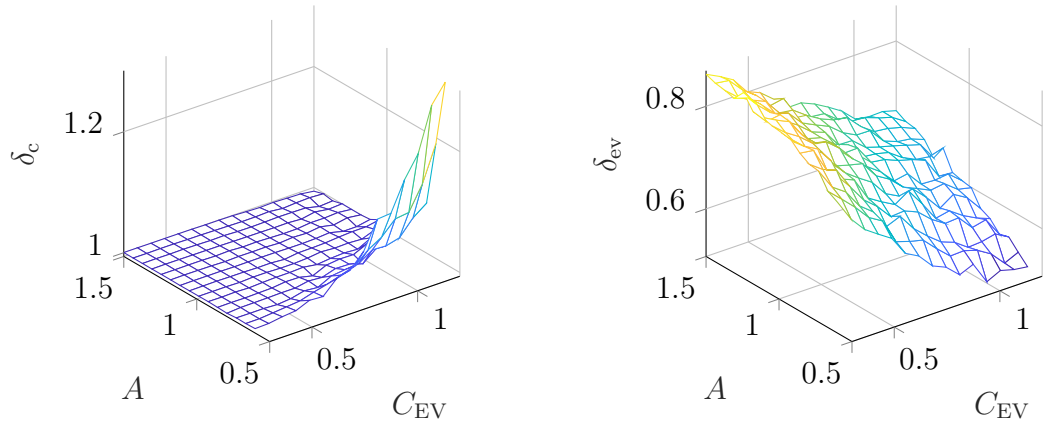
Table 4.1.: Minimum and maximum NMSEs of three spectral parameters for two tested NFT types.

The accuracy for the continuous spectral components does not differ significantly between methods. This could be interpreted as a sign that the truncation threshold T_c was chosen big enough as to not influence the continuous part. The discrete eigenvalue estimation is improved by $\approx 50\%$ for the worst NMSE if $K = 2$ and maybe more interestingly is improved by $\approx 25\%$ for the best achievable NMSE if $K = 2$. Accuracy gains for $K = 3$ are still present, but are reduced in comparison to $K = 2$. Similarly, the improvement for the modulated $b(\lambda_k)$-values is $\approx 20\%$ at best, which, while less pronounced than in the discrete eigenvalue case, is still significant.

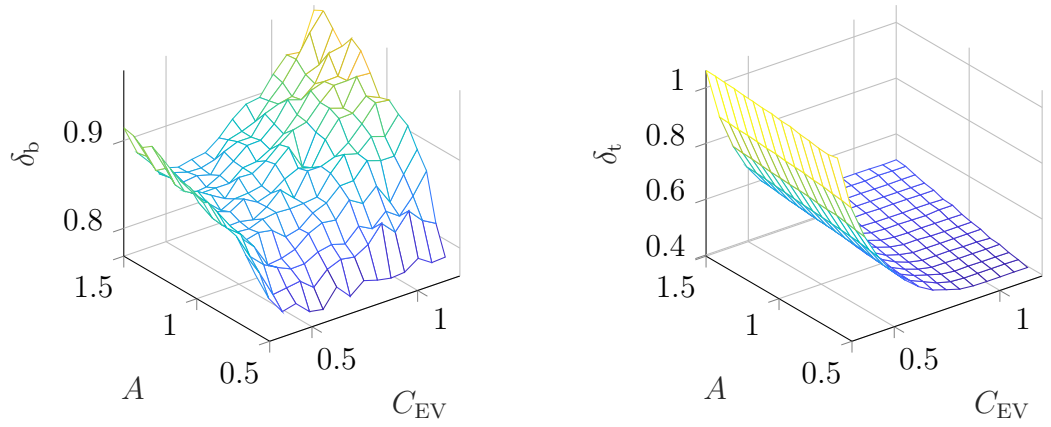
To better visualize the potential improvement of the ER-NFT over the search-based NFT, the ratio between the NMSEs for both methods is computed according to $\delta_\mathrm{x} = e_\mathrm{ER}(\cdot)/e_\mathrm{NR}(\cdot)$ with $\mathrm{x} \in \{\mathrm{c}, \mathrm{ev}, \mathrm{b}\}$. Suffixes c, ev and b denote the ratio for the continuous spectrum, discrete eigenvalues and $b(\lambda_k)$-values, respectively. Additionally, the average execution time was measured and is given as the ratio $\delta_\mathrm{t} = t_\mathrm{ER}/t_\mathrm{NR}$, where t_ER and t_NR are the average execution times for ER-NFT and search-based NFT respectively. The execution time of an algorithm can vary due to many factors and does not always accurately represent the methods complexity. However, the measurements were conducted on the same hardware under similar conditions and thus at least can be used to compare performances for the specific cases tested.

Results for $K = 2$ and varying parameters A and C_EV are given in Fig. 4.16. It can be seen in Fig 4.16a, that in the tested case the detection accuracy for the continuous nonlinear Fourier spectrum did not improve and in some extreme cases the ER-NFT even produced worse results. For the detection of discrete eigenvalues and their corresponding $b(\lambda_k)$-values, the ER-NFT always improved the detection accuracy as depicted in Figs. 4.16b and 4.16c. The algorithm also improved in terms of execution time for almost all tested cases as depicted in Fig. 4.16d. The same evaluation was done for $K = 3$ and the results are given in appendix D.3.

Additionally, data transmission over Raman-amplified fiber was simulated, comparing detection with the search-based NFT and the ER-NFT. The fiber was assumed to be a SSMF with $\alpha = 0.2\,\mathrm{dB/km}$, $\beta_2 = -21.683\,\mathrm{ps^2/km}$ and $\gamma = 1.3\,\mathrm{W^{-1}\,km^{-1}}$. Raman-amplification was implemented with the co- and counterpropagating pump scheme, described in detail in [118]. The total transmission distance was chosen to be 480 km with a span-length of $\mathcal{L}_\mathrm{span} = 80\,\mathrm{km}$. The continuous spectrum was modulated, using 4-PSK constellation points. For the discrete spectrum the $b(\lambda_k)$-values of the discrete eigenvalues were also 4-PSK modulated. The discrete eigenvalues were chosen identically to the back-to-back simulations with $C_\mathrm{EV} = 1$. The pulses were



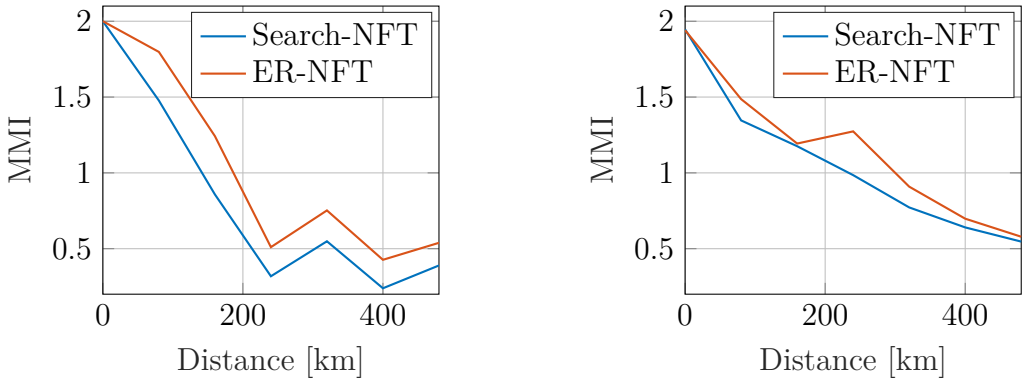
(a) Continuous spectrum NMSE quotient δ_c. (b) Discrete eigenvalues NMSE quotient δ_{ev}.



(c) Discrete b-values NMSE quotient δ_b. (d) Computation time quotient δ_t.

Figure 4.16.: Comparison of detection algorithms for $K = 2$ eigenvalues.

generated with a normalized temporal support of $T_p = 12$ with 1024 samples per pulse and $p = 1 - 10^{-3}$ and were transmitted with the symbol rates given in Fig. 4.17a and 4.17b. Pulses were transmitted independently, assuming enough guard time between pulses to ensure decoupled propagation. Each constellation point was transmitted 100 times for each continuous channel and each discrete eigenvalue. To evaluate the performance of the NFTs, the MMI according to [48] was computed after each fiber span. The values presented in Fig. 4.17 are the mean of the MMIs that were computed separately for the continuous spectrum and each discrete eigenvalue. Note, that the unsteady behavior of the presented curves could be attributed to the breathing motion of the discrete spectrum part of the pulse, which depending on the propagated distance can result in an unfavorable temporal distribution of the discrete spectrum parts



(a) $K = 2$, $R_{\text{Baud}} = 1\,\text{GBaud}$ (b) $K = 3$, $R_{\text{Baud}} = 0.57\,\text{GBaud}$

Figure 4.17.: MI for transmission systems with different NFT algorithms and Raman amplified fiber.

energy. From the results presented in this section, it can be seen that the benefit of utilizing the ER-NFT strongly depends on the parameter range used for modulation. Starting with the discussion of the effect on the accuracy of the continuous spectrum NFT, it can be seen from Fig. 4.16a, that for almost the whole range of value pairs tested, the ER-NFT is not beneficial compared to the search-based approach. For large C_{EV} and low A, there even is a decrease in accuracy from using the ER-NFT. The results imply, that in this case the originally received pulse $q(t)$ should preferably be used for the continuous NFT (compare Figure 4.14). This would increase the computational complexity of the continuous NFT from $\mathcal{O}(C_{\text{t}}(\lambda_K)N^2)$ to $\mathcal{O}(N^2)$, since the received pulse is not truncated prior to the continuous NFT step. On the other hand, this decouples the continuous NFT from the ER step, enabling parallelization, which might again improve computation times on parallel computation architectures. For the discrete eigenvalues and corresponding $b(\lambda_k)$-values, the ER-NFT improves the precision for all tested parameter pairs (see Figs. 4.16b, 4.16c, D.8b, D.8c). For $K = 2$, the error for high C_{EV} and small A values can be reduced by up to $\approx 50\%$, while for small C_{EV} and big A values there is still an improvement of $\approx 14\%$. For the $b(\lambda_k)$-values, there seems to be an optimum value for C_{EV} in the tested region, while the improvement increases for small A. Further, since T_{c} limits the reduction of numerical complexity, the current implementations complexity will be close to the complexity of the search-based method if T_{c} is approximately the temporal width of the received pulse.

Regarding the reduction of computation time, it can be seen that as long as C_{EV}

is large enough, so that the algorithm can sufficiently truncate the pulse in each step, the ER-NFT is able to reduce the computation time. It can also be seen, that the parameter A does not influence the computation time in the tested region. For $K = 2$, a maximum reduction by over 60% can be achieved (see Fig. 4.16d), while for $K = 3$, still a maximum reduction of $\approx 50\%$ is visible (see Fig. D.8d). We also note, that the code for the implementation of the ER-NFT was less optimized than the code of the search-based approach. This could mean that the running time could be improved further by code optimization. In Fig. 4.17 it can be seen, that the achievable rates are improving, if the ER-NFT is used instead of the search-based NFT. Since the ER-NFT does not significantly improve the continuous spectrum detection, the gain stems from the improved detection of constellation points transmitted in the discrete spectrum.

Several other parameters influence the accuracy and computation time of the presented algorithms. In the presented simulations, purely imaginary eigenvalues were chosen. This results in corresponding parts of the pulses, which do not move in their allocated time slot during propagation. This changes if the discrete eigenvalues have some nonzero real component. Parts of the pulse corresponding to these eigenvalues might move towards the edges of the time window and thus might restrict the possible amount of truncation. Another important factor is the spreading of the signal part corresponding to the continuous spectrum. If, due to higher transmission distances, this part of the pulse is spreading, the truncation threshold T_c has to be adapted, which might also restrict the amount of truncation possible. The temporal width of the signal part representing the continuous spectrum is depending on the exact way the continuous spectrum is modulated and thus certain parameter sets might also increase the threshold parameter T_c, reducing the amount of truncation that can be achieved. Note, that contrary to e.g. WDM transmission a separate compensation of linear effects at the receiver to reduce the temporal spreading, or the use of technologies like dispersion compensating fiber (DCF) is not purposeful. Such measures would lead to a removal of effects that are included in the channel model for which the NFT is derived and thus would lead to distortions in the nonlinear Fourier domain. Additionally, compared to the widely used search-based approach, the algorithmic structure of the ER-NFT is less suited for implementation on parallel computing architectures and thus might have drawbacks compared to search-based methods if such architectures are available in the respective system. All these effects have to be considered, regarding the utilization of the presented algorithm in a certain system.

4.4.4. Coefficient Recovery from Nonlinear Spectra

In order to reconstruct a time domain signal from nonlinear spectrum data, some algorithms require the nonlinear Fourier coefficients $a(\lambda)$, $b(\lambda)$ as direct input parameters. This is especially true for some purely continuous spectrum INFTs, as e.g. presented in section 4.4.5. The process to recover the nonlinear Fourier coefficients is depending on the dispersion domain of the fiber channel. The two processes will be detailed for continuous spectrum pulses in the following.

For the defocusing fiber channel ($\mathrm{sgn}(\beta_2) = 1 \rightarrow D < 0$) the nonlinear Fourier coefficients are recovered by using the relation [24]

$$|a(\lambda)|^2 - |b(\lambda)|^2 = 1,\ \lambda \in \mathbb{R}, \tag{4.132}$$

to obtain the absolute value of the first coefficient

$$|a(\lambda)| = \frac{1}{\sqrt{1 - |q_\mathrm{c}(\lambda)|^2}}. \tag{4.133}$$

While this step can in general be done for both, the focusing and the defocusing case, the phase $\angle a(\lambda)$ and the magnitude $|a(\lambda)|$ are only linked by

$$\angle a(\lambda) = \mathcal{H}\{\log(|a(\lambda)|)\}, \tag{4.134}$$

if β_2 is in the defocusing regime.

For the focusing case ($\mathrm{sgn}(\beta_2) = -1 \rightarrow D > 0$), the Riemann-Hilbert (RH) factorization problem, given in [43] has to be solved. Since it was assumed that the generated time-domain pulses have purely continuous spectrum, the RH problem given in [43] is simplified to

$$\boldsymbol{V}^1(t,\lambda) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \frac{1}{2\pi \mathrm{j}} \int_{-\infty}^{\infty} \frac{q_\mathrm{c}^*(\zeta)\mathrm{e}^{-2\mathrm{j}\zeta t}\tilde{\boldsymbol{V}}^1(t,\zeta^*)}{\zeta - (\lambda + \mathrm{j}\epsilon)}\mathrm{d}\zeta, \tag{4.135}$$

$$\tilde{\boldsymbol{V}}^1(t,\lambda^*) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{1}{2\pi \mathrm{j}} \int_{-\infty}^{\infty} \frac{q_\mathrm{c}(\zeta)\mathrm{e}^{2\mathrm{j}\zeta t}\boldsymbol{V}^1(t,\zeta)}{\zeta - (\lambda - \mathrm{j}\epsilon)}\mathrm{d}\zeta, \tag{4.136}$$

where ϵ is some small positive value and vectors

$$\boldsymbol{V}^1(t,\lambda) = \boldsymbol{\psi}(t,\lambda)\mathrm{e}^{-\mathrm{j}\lambda t},\ \tilde{\boldsymbol{V}}^1(t,\lambda^*) = \tilde{\boldsymbol{\psi}}(t,\lambda^*)\mathrm{e}^{\mathrm{j}\lambda t}, \tag{4.137}$$

$$\boldsymbol{V}^2(t,\lambda) = \boldsymbol{\phi}(t,\lambda)\mathrm{e}^{\mathrm{j}\lambda t},\ \tilde{\boldsymbol{V}}^2(t,\lambda^*) = \tilde{\boldsymbol{\phi}}(t,\lambda^*)\mathrm{e}^{-\mathrm{j}\lambda t}, \tag{4.138}$$

result in boundary conditions

$$\boldsymbol{V}^1(t,\lambda) \to \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ for } t \to \infty, \tag{4.139}$$

$$\tilde{\boldsymbol{V}}^1(t,\lambda^*) \to \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ for } t \to \infty, \tag{4.140}$$

$$\boldsymbol{V}^2(t,\lambda) \to \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ for } t \to -\infty, \tag{4.141}$$

$$\tilde{\boldsymbol{V}}^2(t,\lambda^*) \to \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ for } t \to -\infty. \tag{4.142}$$

The projection equation (4.28) , by which the nonlinear Fourier coefficients are defined, is now altered to be [43]

$$\boldsymbol{V}^2(t,\lambda) = a(\lambda)\tilde{\boldsymbol{V}}^1(t,\lambda^*) + b(\lambda)\mathrm{e}^{2\mathrm{j}\lambda t}\boldsymbol{V}^1(t,\lambda). \tag{4.143}$$

Since the nonlinear Fourier coefficients do not depend on time, parameter t can be chosen freely. To simplify the computation, $t \to -\infty$ is chosen. As a result Eq. (4.143) can be used to state

$$a(\lambda) = \frac{V_2^1(-\infty,\lambda)}{\tilde{V}_1^1(-\infty,\lambda^*)V_2^1(-\infty,\lambda) - \tilde{V}_2^1(-\infty,\lambda^*)V_1^1(-\infty,\lambda)}, \tag{4.144}$$

where $V_1^1(t,\lambda)$, $V_2^1(t,\lambda)$ and $\tilde{V}_1^1(t,\lambda^*)$, $\tilde{V}_2^1(t,\lambda^*)$ are the first and second entry of $\boldsymbol{V}^1(t,\lambda)$ and $\tilde{\boldsymbol{V}}^1(t,\lambda^*)$ respectively. The second coefficient is obtained by using the relation from Eq. 4.35.

The RH problem stated above can also be formulated for purely discrete spectrum pulses by using the system of equations given in [43] and simplifying it accordingly. The problem can also be solved without simplifications for the full nonlinear spectrum. We restrict ourselves to the treatment above, since for all other cases, more effective methods will be presented in sections 4.4.6 and 4.4.7. A numerical method to solve

the RH problem and obtain vectors $\boldsymbol{V}^1(t,\lambda)$, $\tilde{\boldsymbol{V}}^1(t,\lambda^*)$ will be given in section 4.4.5.

4.4.5. Inverse NFT for the Continuous Spectrum

The INFT for purely continuous spectrum pulses aims to reconstruct the time domain signal $q(t)$ from, either the continuous spectrum function $q_\mathrm{c}(\lambda)$, or the corresponding nonlinear Fourier coefficients $a(\lambda)$, $b(\lambda)$. This can be achieved by several methods such using a Gelfand-Levitan-Marchenko (GLM) [119], [120] or RH solver [28], or by employing inverse methods directly derived from some well known forward transformation algorithms, such as the inverse AL-discrete layer peeling (DLP) method [23], [121]. The RH-solver and the inverse AL-DLP method will be shortly reviewed in the following.

Riemann-Hilbert Solver

The RH-solver for purely continuous spectrum pulses utilizes the discretized versions of Eqs. (4.135) and (4.136), which can be rearranged as

$$\underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\boldsymbol{B}_1} = \tilde{\boldsymbol{V}}^1(t,\lambda[m]) - \frac{1}{2\pi\mathrm{j}} \sum_{n=1}^{N} \underbrace{\frac{q_\mathrm{c}[m]\mathrm{e}^{2\mathrm{j}\lambda[n]t}\Delta\lambda}{\lambda[n]-(\lambda[m]-\mathrm{j}\epsilon)}}_{C_{mn}} \boldsymbol{V}^1(t,\lambda[n]), \tag{4.145}$$

$$\underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{\boldsymbol{B}_2} = \boldsymbol{V}^1(t,\lambda[m]) - \frac{1}{2\pi\mathrm{j}} \sum_{n=1}^{N} \underbrace{\frac{q_\mathrm{c}^*[m]\mathrm{e}^{-2\mathrm{j}\lambda[n]t}\Delta\lambda}{\lambda[n]-(\lambda[m]+\mathrm{j}\epsilon)}}_{C_{mn}^*} \tilde{\boldsymbol{V}}^1(t,\lambda[n]), \tag{4.146}$$

where $\Delta\lambda = \lambda[n+1]-\lambda[n]$ is the sample-width. The discretized system can be solved by setting up the matrix equation $\boldsymbol{\mathcal{B}} = \boldsymbol{A}\boldsymbol{\mathcal{V}}$, where $\boldsymbol{\mathcal{B}}$ contains vectors $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$, $\boldsymbol{\mathcal{V}} = [\boldsymbol{\mathcal{V}}^1, \tilde{\boldsymbol{\mathcal{V}}}^1] = [\boldsymbol{V}^1(t,\lambda[1]),\cdots,\boldsymbol{V}^1(t,\lambda[N]),\tilde{\boldsymbol{V}}^1(t,\lambda[1]),\cdots,\tilde{\boldsymbol{V}}^1(t,\lambda[N])]$ and matrix

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I}_N & -\frac{1}{2\pi\mathrm{j}}\boldsymbol{C}^* \\ -\frac{1}{2\pi\mathrm{j}}\boldsymbol{C} & \boldsymbol{I}_N \end{bmatrix}, \tag{4.147}$$

where $\boldsymbol{C}$ has entries C_{mn}. It can be seen that, in order to achieve a certain accuracy the necessary matrix inversion of $\tilde{\boldsymbol{A}} = \boldsymbol{A}^{-1}$ can become quite computationally intense.

The time-domain pulse can then be recovered approximately by using [43]

$$q^*[m] \approx \frac{1}{\pi} \sum_{n=1}^{N} q_\mathrm{c}[n] \mathrm{e}^{2\mathrm{j}\lambda[n]t[m]} V_2^1(t[m], \lambda[n]) \Delta\lambda. \tag{4.148}$$

Note, that the derivation of this approximation is rather involved and can be found in [43, Ch. VII B], [43, App. E] and [43, App F]. It can be seen, that for every recovered time-domain sample, the matrix given in 4.147 has to be inverted. Thus, the complexity of this method is quite high and other algorithms, such as the AL-DLP method described below, reduce the complexity of the transformation, while maintaining or even improving the accuracy. There also is another way to determine $\boldsymbol{\mathcal{V}}$, which is described in appendix B.1.

Inverse Ablowitz-Ladik Discrete Layer Peeling

The inverse Ablowitz-Ladik discrete layer peeling (AL-DLP) algorithm is another way to recover a time-domain signal that only has a continuous spectrum from its nonlinear Fourier coefficients. Once more, the transformation is executed for some constant propagation distance z and thus this parameter is omitted in the following. Additionally, the parameter $z(\lambda) = e^{-2\mathrm{j}\lambda h}$ is defined and should not be confused with the propagation parameter. In the following, quantities $a(t[n],\, \lambda[\ell]) = a[n,\, \ell]$, $b(t[n],\, \lambda[\ell]) = b[n,\, \ell]$ are quantities, that contain all the solutions of the iterative steps from Eq. 4.84 in section 4.4.1, that are done to obtain the nonlinear Fourier coefficients $a[N_\mathrm{s} - 1,\, \ell]$, $b[N_\mathrm{s} - 1,\, \ell]$ on the discretized real-line. The number of discretization steps N_s is equal in time and continuous nonlinear spectrum. Indices are bound by $0 \leq n \leq N_\mathrm{s} - 1$ and $0 \leq \ell \leq N_\mathrm{s} - 1$. Note, that this method requires $N_\mathrm{s} h \Delta\lambda = \pi$ to be fulfilled, where $\Delta\lambda$ is the sample-width on the real-line λ [23].

In a first step, using the nonlinear Fourier coefficients on the real-line $a[N_\mathrm{s} - 1, \ell]$, $b[N_\mathrm{s} - 1, \ell]$, the quantities

$$A[N_\mathrm{s} - 1, \ell] = a[N_\mathrm{s} - 1, \ell], \tag{4.149}$$

$$B[N_\mathrm{s} - 1, \ell] = z^{-(n_0 + N_\mathrm{s} - 1) + \frac{1}{2}} b[N_\mathrm{s} - 1, \ell], \tag{4.150}$$

$$\tag{4.151}$$

with $n_0 = \frac{t_1}{h}$ are computed. Using these coefficients, the vector

$$\hat{\boldsymbol{A}}[n] = \left(A[n, \mathrm{e}^{-\mathrm{j}2\lambda[\ell=0]t[n]}], \cdots, A[n, \mathrm{e}^{-\mathrm{j}2\lambda[\ell=N_\mathrm{s}-1]t[n]}]\right), \tag{4.152}$$

can be computed and its analog $\hat{\boldsymbol{B}}[n]$ can be computed similarly. Note that, while the equation above is given for general values n, at first only the vectors for $n = N_\mathrm{s} - 1$ can be obtained.

The iterative part of the algorithm is executed in a domain obtained by using the discrete Fourier transform (DFT) on quantities $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$. More exactly, the relation given in [23] is

$$\boldsymbol{A}[n] = \frac{1}{N_\mathrm{s}} \boldsymbol{e} \circ \mathrm{DFT}\{\hat{\boldsymbol{A}}[n]\} \tag{4.153}$$

where the relation for $\boldsymbol{B}[n]$ and $\hat{\boldsymbol{B}}[n]$ is obtained by exchanging $\boldsymbol{A}[n]$ for $\boldsymbol{B}[n]$ and $\hat{\boldsymbol{A}}[n]$ for $\hat{\boldsymbol{B}}[n]$. Operation $\circ$ is the element-wise multiplication of two vectors and $\boldsymbol{e} = (\mathrm{e}^{-\mathrm{j}2l_1h\cdot 0}, \mathrm{e}^{-\mathrm{j}2l_1h\cdot 1}, \cdots, \mathrm{e}^{-\mathrm{j}2l_1h\cdot(N_\mathrm{s}-1)})$. Further $\lambda \in [l_1, l_2]$ and thus l_1 is the smallest value of the nonlinear temporal support. Due to the way they are obtained, parameters $\boldsymbol{A}[n] = (A_0[n], \cdots, A_{N_\mathrm{s}-1}[n])$ and $\boldsymbol{B}[n] = (B_0[n], \cdots, B_{N_\mathrm{s}-1}[n])$ are said to be in the *nonlinear time domain*, even though this does not have any instantly recognizable physical significance. In [23] an iterative scheme to obtain the missing quantities $\boldsymbol{A}$, $\boldsymbol{B}$ for $n < N_\mathrm{s} - 1$ is developed. The relations for iterating from $n = N_\mathrm{s} - 1$ towards $n = 0$ are given by [23]

$$\boldsymbol{A}[n] = c_n \cdot (\boldsymbol{A}[n+1] - Q[n]\boldsymbol{B}[n+1]), \tag{4.154}$$

$$\boldsymbol{B}[n] = c_n \cdot \mathrm{shift}\,(Q^*[n]\boldsymbol{A}[n+1] + \boldsymbol{B}[n+1]). \tag{4.155}$$

Here $c_n = \frac{1}{\sqrt{1+|Q[n]|^2}}$ and the shift-operation is defined as $\mathrm{shift}(\boldsymbol{x}) = (x_2, x_3, \cdots, x_N, 0)$, where $\boldsymbol{x} = (x_1, x_2, \cdots, x_N)$. The time-domain signal is obtained by relation [23]

$$Q[n] = \frac{B_0[n+1]}{A_0[n+1]}, \tag{4.156}$$

where the $Q[n] = q[n]/h$ as first defined in section 4.4.1.

It can be seen, that when starting with $\boldsymbol{A}[N_\mathrm{s} - 1]$ and $\boldsymbol{B}[N_\mathrm{s} - 1]$, which can be obtained from the nonlinear Fourier coefficients, the first sample of $Q[n]$ can already be computed. This value then can be used in the next iteration of Eqs. (4.154) and (4.155), which in return enable the computation of the next sample of $Q[n]$. Thus, an

iterative algorithm to obtain the discretized time-domain pulse $q[n]$ from the nonlinear Fourier coefficients $a(\lambda)$, $b(\lambda)$ is formed. This algorithm has proven to be numerically advantageous over directly solving the discretized RH-system in many scenarios.

4.4.6. Inverse NFT for the Discrete Spectrum

The discrete spectrum inverse NFT can also be implemented by solving a RH-problem. Assuming that, there are discrete eigenvalues in the upper complex halfplane and that $q_\mathrm{c}(\lambda) = 0$ everywhere on the real line, the simplifications of the RH-system in [43] done in Eqs. (4.135) and (4.136) have to be altered accordingly. The resulting equations are

$$\boldsymbol{V}^1(t,\lambda_m) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \sum_{k=1}^{K} \frac{q_\mathrm{d}^*(\lambda_k^*)\mathrm{e}^{-2\mathrm{j}\lambda_k^* t}\tilde{\boldsymbol{V}}^1(t,\lambda_k^*)}{\lambda_m - \lambda_k^*} \tag{4.157}$$

$$\tilde{\boldsymbol{V}}^1(t,\lambda_m^*) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \sum_{k=1}^{K} \frac{q_\mathrm{d}(\lambda_k)\mathrm{e}^{2\mathrm{j}\lambda_k t}\boldsymbol{V}^1(t,\lambda_k)}{\lambda_m^* - \lambda_k} \tag{4.158}$$

Then the method outlined in section 4.4.5 can be used to compute the time-domain pulse corresponding to the discrete spectrum data. It can be seen, that for a realistic number of discrete eigenvalues, the complexity is not as big of an issue as it might be for the continuous spectrum RH-solver. However, a more commonly used method, namely the Darboux transform (DT) method, is very accurate and in many cases exhibits a comparable or lower computational complexity than the RH-solver. The inverse NFT via DT [28] is outlined in the following.

The main advantage of the DT over e.g. a discrete spectrum RH scheme is, that it does not introduce numerical errors due to discretization and does not involve matrix inversions that potentially can be ill-conditioned. The DT as presented here, can only generate pulses from purely discrete nonlinear spectra. However, a modified version of the DT, that can deal with the full nonlinear spectrum, will be presented in section 4.4.7.

The addition of a discrete eigenvalue λ_k to the nonlinear discrete spectrum of a time-domain signal $q(t)$ is facilitated by the *auto-Bäcklund transformation* [122]

$$q^{(k+1)}(t) = q^{(k)}(t) + 2\mathrm{j}(\lambda_k^* - \lambda_k)\frac{v_1(t,\lambda_k)v_2^*(t,\lambda_k^*)}{|v_1(t,\lambda_k)|^2 + |v_2(t,\lambda_k)|^2}, \tag{4.159}$$

where $q^{(k+1)}(t)$ is the time domain pulse, for which the discrete nonlinear spectrum was augmented to include λ_k. The parameters $v_1(t, \lambda_k)$, $v_2(t, \lambda_k)$ are the entries of vectors $v(t, \lambda_k)$, which was obtained for $q^{(k)}(t)$. Note, that Eq. (4.159) is identical to Eq. (4.131) in section 4.4.3 for the eigenvalue removal method [113]. For the first update according to Eq. (4.159), it is assumed that the continuous spectrum of the current pulse is zero everywhere and that there are no discrete eigenvalues present yet. Thus, the corresponding time domain pulse has to be the all-zero pulse ($q^{(1)}(t) = 0$). For this pulse the vectors $v(t, \lambda_k)$ can be computed analytically according to [106]

$$v_1(t, \lambda_k) = \mathrm{e}^{-\mathrm{j}\lambda_k t} \tag{4.160}$$

$$v_2(t, \lambda_k) = -\frac{q_\mathrm{d}(\lambda_k)}{\lambda_k - \lambda_k^*} \prod_{i=1, i\neq k}^{K} \frac{\lambda_k - \lambda_i}{\lambda_k - \lambda_i^*} \mathrm{e}^{\mathrm{j}\lambda_k t}. \tag{4.161}$$

After the update step, the vectors $\boldsymbol{v}(t, \lambda_k)$ are no longer valid for the new pulse shape $q^{(k+1)}(t)$ and thus they have to be updated according to [122, Theorem 1]

$$\boldsymbol{v}(t, \lambda_{k+1}) = (\lambda_k \boldsymbol{I} - \boldsymbol{\Sigma})\boldsymbol{v}(t, \lambda_k) \tag{4.162}$$

with $\boldsymbol{\Sigma} = \boldsymbol{S}\boldsymbol{\Gamma}\boldsymbol{S}^{-1}$, $\boldsymbol{S} = [\boldsymbol{v}(t, \lambda_k), \tilde{\boldsymbol{v}}(t, , \lambda_k^*)]$ and $\boldsymbol{\Gamma} = \mathrm{diag}(\lambda_k, \lambda_k^*)$.

It can be seen, that Eqs. (4.159) and (4.162) result in a method, that can be used to add discrete eigenvalues to an initial trivial pulse. The iterative nature of this process is depicted Fig. 4.18. It can be seen, that for some initial time domain

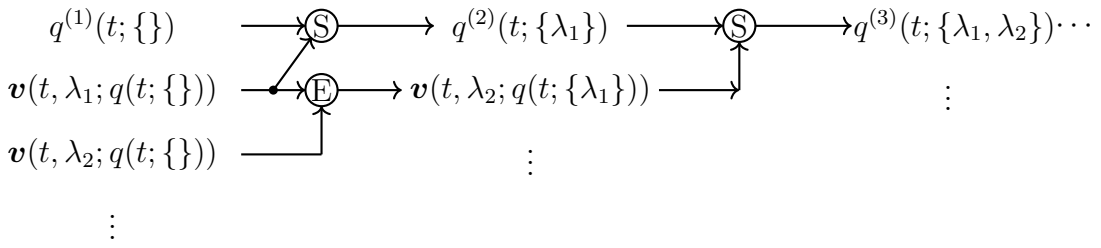
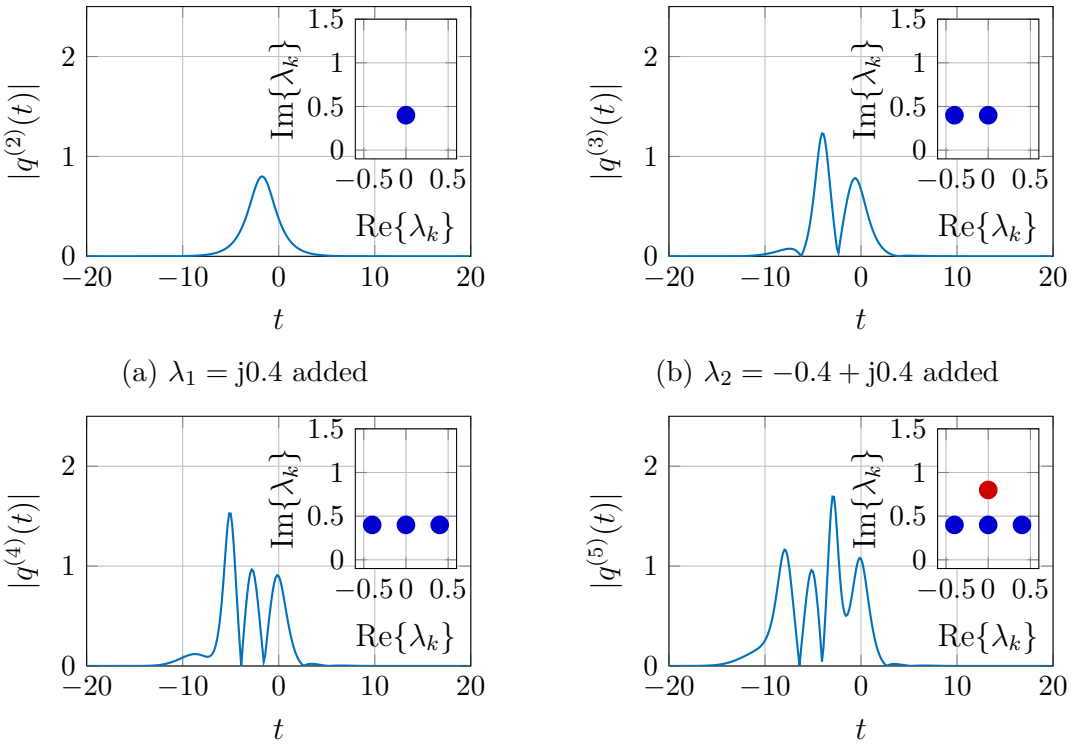


Figure 4.18.: Diagram of Darboux-Transform iterations (adapted from [28])

signal $q(t;\{\})$, where $\{\}$ is the empty discrete spectrum and corresponding eigenvectors $\boldsymbol{v}(t, \lambda_i, q(t;\{\}))$, several discrete eigenvalues can be added to the discrete spectrum of the pulse.

A depiction of the (intermediate) time domain pulses and their corresponding set of discrete eigenvalues, during the generation of a $K = 4$ soliton, is given in Fig. 4.19,

negelecting the trivial zero-pulse at the beginning. It can be seen, that after the first



(a) $\lambda_1 = \mathrm{j}0.4$ added

(b) $\lambda_2 = -0.4 + \mathrm{j}0.4$ added

(c) $\lambda_3 = 0.4 + \mathrm{j}0.4$ added

(d) $\lambda_4 = \mathrm{j}0.8$ added

Figure 4.19.: Iterative soliton generation using the Darboux Method ($K = 4$)

step the distinctive hyperbolic secant shape of a fundamental soliton is obtained. A faster implementation of the DT algorithm can be found in [106].

While no conclusive simulations were conducted, to compare the RH-solver method with the DT, the difference in the results between both methods seems to be rather small for small numbers of discrete eigenvalues. With an increase in K, the error generated by the RH-solver method increases as well, up to the point where the results get unstable. The onset of this instability is depicted for $K = 8$ in Fig. 4.20.

4.4.7. Full Nonlinear Spectrum INFT

Similar to the INFT for continuous and discrete spectrum, the time domain pulse for a general nonlinear spectrum, consisting of non-zero continuous and discrete parts, can be computed using the RH-solver approach. To do so the steps for the previous cases have to be done without simplifications of the initial RH-system. This approach was

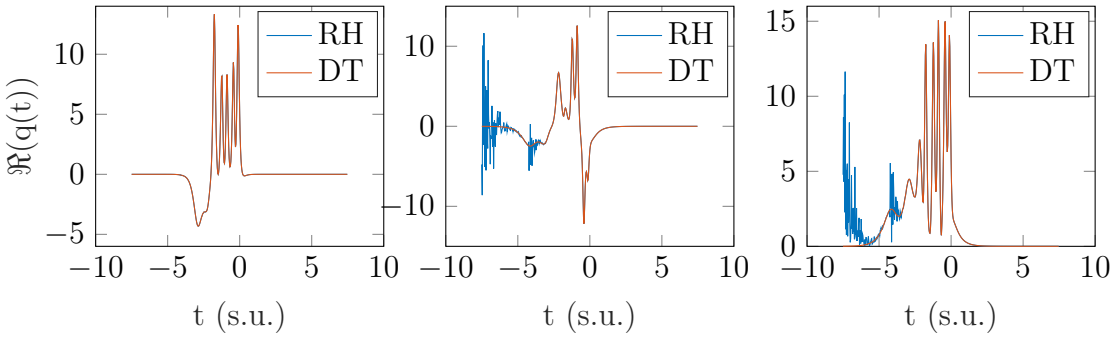


Figure 4.20.: Unstable behaviour of RH-solver for $K = 8$, $N = 256$.

not implemented. Instead, the DT algorithm presented in section 4.4.6 was augmented, to accept an initial time-domain solution other than the trivial zero-pulse $q_{\text{seed}}(t) = 0 \,\forall\, t$ [123], [124].

In the following, this approach will be called the *seeded Darboux transform*. The steps of this inverse transformation are layed out in Fig. 4.21.

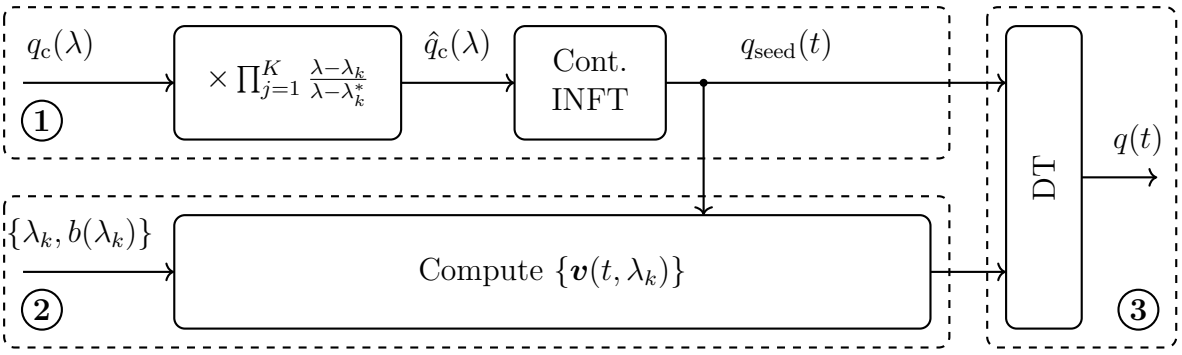


Figure 4.21.: Block diagram of seeded DT (based on [124])

If, in the i-th iteration of the DT, a discrete eigenvalue λ_i is added to the nonlinear discrete spectrum of a pulse, its continuous spectrum $q_c(\lambda)$ is altered according to [106]

$$q_c(\lambda)^{(i+1)} = \frac{\lambda - \lambda_i^*}{\lambda - \lambda_i} q_c(\lambda)^{(i)}. \tag{4.163}$$

Thus, in part ① of Fig. 4.21, the targeted nonlinear continuous spectrum $q_c(\lambda)$ is predistorted before the continuous INFT is used to compute $q_{\text{seed}}(t)$. The continuous INFT was implemented by the inverse AL-DLP scheme [23], but in general any continuous INFT method can be used.

Since the initial pulse for the DT in part ③ is no longer $q_{\text{seed}}(t) = 0 \,\forall\, t$, the functions $\{\boldsymbol{v}(t, \lambda_k)\}$ are no longer known analytically. To be able to add discrete eigenvalues λ_k to the nonlinear discrete spectrum of $q_{\text{seed}}(t)$, the correct values for $\{\boldsymbol{v}(t, \lambda_k)\}$ have to be computed, using the prescribed nonlinear discrete spectral components $\{\lambda_k, b(\lambda_k)\}$ and the initial pulse $q_{\text{seed}}(t)$. This computation, in part ② of Fig. 4.21, can be done by slightly modifying any of the algorithms presented in section 4.4.1 (except for the LP-algorithm), to output the set of vectors $\{\boldsymbol{v}(t, \lambda_k)\}$ directly.

In Fig. 4.22, the method outlined above was used to construct a time-domain rectangular pulse $q(t)$ with two discrete eigenvalues and non-zero nonlinear continuous spectrum. The resulting pulse $q(t)$ should be similar to the analytical result $q_{\text{true}}(t)$.

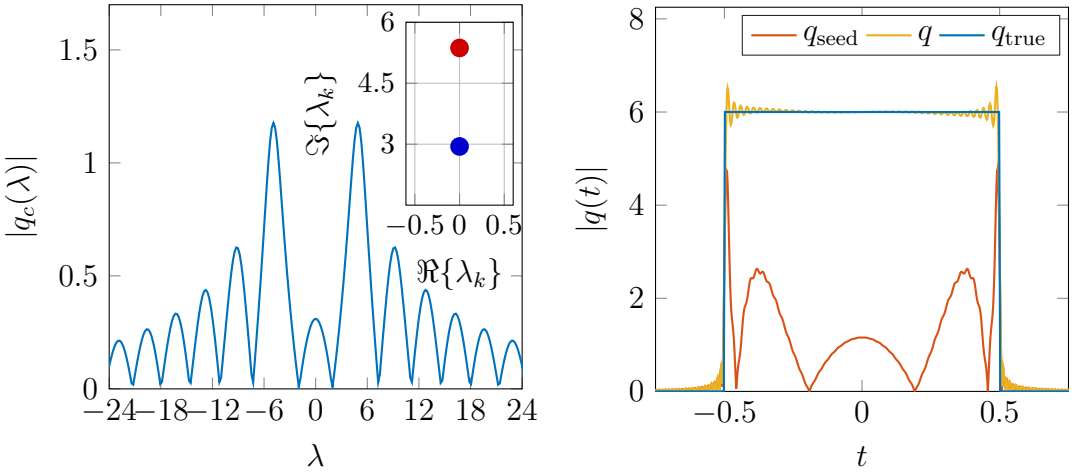


Figure 4.22.: Reconstruction of rectangular pulse via seeded DT ($A = 6$, $T = 1$)

It can be seen, that apart from some jitter at the edges of the rectangular pulse, the result is quite accurate even for this rather problematic pulse-shape. Before we conclude this section, we would like to point out that using a RH-solver or seeded DT approach are not the only ways to implement an INFT for continuous and discrete spectrum pulses. Some other methods can be found in [92], [125]–[127].

4.5. On the Time-Bandwidth Product of Solitons

The TBP for the fundamental soliton, is known analytically and depends on the amount of in-window energy, as described in section 4.3.2. However, this perspective is only of limited practical use, when modulation is added to the picture. In the

following, some closer observations for some specific systems are made, to gain further insight regarding the requirements for soliton transmission. To make a clear distinction between the TBP of a single pulse and the bandwidth and time-slot requirements of some modulated pulse-train, the expression time-bandwidth requirement (TBR) is introduced for the latter. It is defined, as the product of the minimum time-slot per pulse and bandwidth per channel, that has to be allocated, to ensure some amount of in-window energy is supported. For the TBR, it is always assumed that the soliton-train is propagating ideally in the channel and no guard-times or guard-bands are introduced to deal with perturbations or residual ISI. Note, that in realistic system setups, guard-times are beneficial to mitigate interference between neighboring signals. This is even true for purely solitonic systems, since the truncation of solitonic pulses according to some factor p leads to radiative signal components, which interfere with adjacent pulses via dispersion. Guard-bands are necessary in realistic setups as well, to e.g. relax the requirements on (RX)-filters. Also, the NFT framework is not suited to deal with the effects of linear frequency multiplexing on the nonlinear spectrum analytically, thus to approximate valid conditions for the NFT, the crosstalk between linearly multiplexed channels has to be kept rather small. Of course the resources spent on guard-times and guard-bands have to be incorporated into the evaluation of a systems SE. However, since linearly multiplexed NFT modulated channels are not studied in detail in this thesis, the aforementioned effects are not considered in the development of the TBR studied in this section. Similarly, the additional guard-times that might be needed to avoid interactions between temporally adjacent pulses is neglected since there is no known way to analytically describe the temporal broadening of the radiative parts of the signal generated by truncation. Nonetheless, the TBR is a measure that is more geared towards the evaluation of certain transmission system requirements, than the TBP relations derived e.g. in section 4.3.2.

4.5.1. TBR for Multi-Ring Modulated Trains of Fundamental Solitons

If the solitonic pulses in the transmission-train are modulated, such that they are not dislocated, depending on the modulation symbol, then some fixed offset of the pulse-center from zero becomes meaningless and thus it can be assumed that all pulses in the alphabet are zero-centered. As a result, $|q_\mathrm{d}(\lambda_1)| = 2\eta_1$ has to be fulfilled. This restricts the modulation of $q_\mathrm{d}(\lambda_1)$ to a single-ring PSK constellation. Further, the

pulse-position is assumed to be fixed for varying z. Thus, the discrete eigenvalue λ_1 is purely imaginary. If the above assumptions are true, the expressions for pulse duration (Eq. (4.64)) and linear bandwidth (Eq. (4.66)) from section 4.3.2 are identical to the time-slot and bandwidth requirements of the soliton transmission system. As a result the TBP as given in Eq. (4.67) is identical to the TBR of the system. In the following the requirement for this PSK modulated, zero-centered pulse transmission will be denoted by $\mathrm{TBR}_{\mathrm{zc}}$.

While still requiring λ_1 to be purely imaginary, the treatment of the TBR can now be extended to multi-ring PSK modulation. For MPSK, modulation symbols $|q_{\mathrm{d}}(\lambda_1)| \neq 2\eta_1$ can occur, resulting in solitons that are no longer zero-centered. Thus, while the expressions in section 4.3.2 are still valid on a symbol basis, the time-slot for a transmitted pulse has to be increased. Note that, in addition to the phases located on each ring, the different ring amplitudes result in PPM in time-domain and thus the TBR depends on the exact modulation alphabet, more precisely the exact ring amplitudes used. It is thus not straightforward to obtain expressions similar to the ones given in section 4.3.2.

The necessary temporal support can be approximated by assuming the maximum/minimum time values marking the time-slot are shifted by $\ln(|q_{\mathrm{d}}(\lambda_1)|/2\eta)$ (see Eq. (4.62)). This results in the TBR for the MPSK case being well approximated by

$$\begin{aligned}\mathrm{TBR}_{\mathrm{MPSK}} \approx & \mathrm{TBR}_{\mathrm{zc}} + B_p \left(\left| \log \left(\frac{\min(|q_{\mathrm{d}}(\lambda_1)|)}{2\eta} \right) \right| + \left| \log \left(\frac{\max(|q_{\mathrm{d}}(\lambda_1)|)}{2\eta} \right) \right| \right) && (4.164) \\ = & \mathrm{TBR}_{\mathrm{zc}}(1 + f_{\mathrm{ppm}}), && (4.165)\end{aligned}$$

assuming, the constellation point with the smallest absolute value fulfills $\min(|q_{\mathrm{d}}(\lambda_1)|) \leq 2\eta$ and the constellation point with the highest absolute value fullfills $\max(|q_{\mathrm{d}}(\lambda_1)|) \geq 2\eta$. Note that, this choice of maximum and minimum absolute values for $q_{\mathrm{d}}(\lambda_1)$ is not a real restriction since, a fixed temporal offset for all transmission pulses does have no impact on the system and thus substitute systems that fulfill the above restrictions can always be found. In Fig. 4.23, the approximated TBR for MPSK, according to (4.164) is shown alongside a numerical computation of the TBR, which is assumed to be quite exact.

The numerical computation of the TBR is done by generating fundamental solitons for each ring in the MPSK constellation. Since the phase of $q_{\mathrm{d}}(\lambda_1)$ does not influence the pulses position in the allotted time window, any phase can be chosen. The

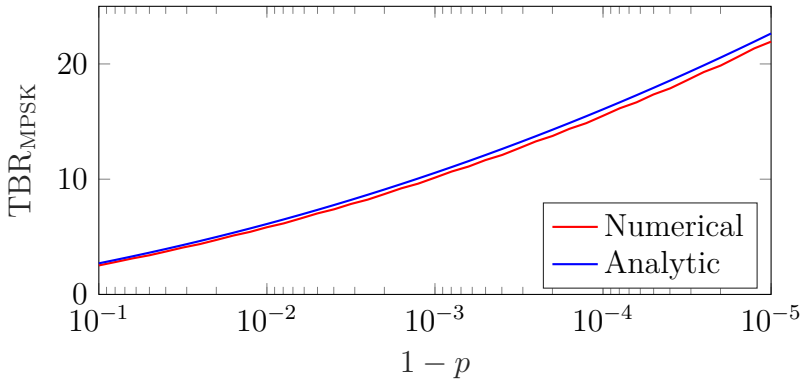


Figure 4.23.: TBR of MPSK modulated fundamental soliton ($|q_{\mathrm{d}}(\lambda_1)| \in \{0.0946,\ 0.7232,\ 5.5307,\ 42.2948\}$). Red: Numerical evaluation, Blue: Closed form approximation

necessary time window and bandwidth for transmission is now measured according to some factor p to obtain the numerical result in Fig. 4.23. Of course the factor p has to be identical for both of the shown curves. It can be seen, that the theoretical approximation slightly overestimates the TBR, but overall fits the real TBR quite well. Note that the results depicted in Fig. 4.23 are obtained for the specific values $|q_{\mathrm{d}}(\lambda_1)|$ given in the caption, but the slight overestimation compared to the measurements did occur for all tested scenarios. While no formal proof will be given, the result can be explained as follows. Given a modulation alphabet, for which pulses are shifted from the center by up to $t_{\mathrm{shift,min}}$ and $t_{\mathrm{shift,max}}$, the time-slot has to be expanded to still fulfill the in-window energy constraint. Eq. (4.164) increases the time-slot by the value $|t_{\mathrm{shift,min}}|+|t_{\mathrm{shift,max}}|$. Since there is some additional in-window energy, due to the pulses tail that is shifted further into the window, the extension of the time-slot could be chosen slightly smaller. Due to the fast-decaying tails of fundamental solitons, this deviation is rather small. Thus Eq. (4.164) approximates the TBR for the MPSK modulated case rather well.

4.5.2. TBR for Multi-Ring Modulated Trains of Multi-Solitons

A general closed form expression for the TBP of solitons with more than one discrete eigenvalue is currently unknown. While some approximations for specific cases exist [116], the TBP has to be evaluated numerically for most cases.

In the following, the discussion is once more limited to cases where all discrete eigenvalues are located on the imaginary axis ($\Re\{\lambda_k\} = 0$ for all $k \in (1, K)$). While

discrete eigenvalues with non-zero real parts are in general usable in e.g. soliton transmission systems [128], they result in the separation of the multi-soliton into several fundamental solitons during propagation. As a result, the TBR additionally becomes dependent on total transmission distance $\mathcal{L}$. For a multi-soliton with purely imaginary eigenvalues, the TBR is confined to some finite value even for $\mathcal{L} \to \infty$ and thus is more suited for the following analysis.

Multi-solitons ($K > 1$) no longer retain their shape during propagation. Instead their envelope changes periodically and while the period might become infinite in some cases, some in-window energy of these pulses is supported on a finite temporal support at any time. Due to the fact, that the temporal and linear spectral width of this pulses changes in z, without dispersing for $z \to \infty$, this type of pulses is often called a *breather-soliton*.

If now the time-slot T_{N-SOL}, dependent on some out-of window energy percentage p for a fixed set of discrete eigenvalues $\{\lambda_k\}$, in a multi-soliton train setup is to be set, it has to fulfill

$$T_{N-SOL} = \max_{\{\phi_k\} \in [0,2\pi)} \mathcal{T}(q(t, \{\lambda_k \; q_\mathrm{d}(\lambda_k)\}), p) \tag{4.166}$$

where $\mathcal{T}(\cdot,\cdot)$ is a function giving the temporal width of a pulse, according to the used in-window energy percentage p. The parameter ϕ_k is the phase of the nonlinear spectral amplitude $q_\mathrm{d}(\lambda_k)$. We note that, independent of the phases per ring modulated at the transmitter, all phase combinations potentially occur due to the λ_k dependent phase-shifts during propagation (see Eq. (4.51)). Thus, to limit ISI to the amount prescribed by p, it is necessary to also evaluate phase values that are not in the modulation alphabet. Note, that in practical systems, $\mathcal{L}$ has some finite value, so especially for short distances the range of the values $\{\phi_k\}$ is potentially smaller in these cases and thus, the time-slot T_{N-SOL} might also be smaller than for the general $\mathcal{L} \to \infty$ case. This implies, that the time-slot can be optimized for each distance and modulation alphabet to increase the SE of the system. Further, according to Eq. (4.51) the propagation of nonlinear spectral amplitudes is given by

$$q_\mathrm{d}(\lambda_k, z) = |q_\mathrm{d}(\lambda_k, z = 0)| \mathrm{e}^{-\mathrm{j}(\phi_k(z=0) - 4\eta^2 z)}. \tag{4.167}$$

It can be seen, that the absolute values of the nonlinear spectral amplitudes are not affected during propagation in z and thus are not used in the maximization in Eq. (4.166). Note, that the time-domain shape still is dependent on the absolute values

$|\lambda_k|$ and thus time-slot T_{N-SOL} is only optimized for this specific set of rings. However, if $\mathcal{L} \to \infty$ is assumed and thus $\{\phi_k\} \in [0, 2\pi)$, the number of phases per ring and their exact distribution on the rings in the modulation alphabet does not influence T_{N-SOL}.

A similar observation can be made for the linear bandwidth, allocated for such multi-soliton pulse-trains and thus an expression of the needed bandwidth can be found by

$$B_{N-SOL} = \max_{\{\phi_k\}\in[0,2\pi)} \mathcal{B}(q(t, \{\lambda_k \; q_\mathrm{d}(\lambda_k)\}), p) \tag{4.168}$$

where $\mathcal{B}(\cdot,\cdot)$ is a function returning the linear bandwidth of the pulse, according to some in-band energy percentage p. Maximization (4.168) can again become prohibitively large and thus is as well assumed to be approximated sufficiently well by testing a randomly generated pulse-set of size N_i.

To numerically evaluate Eqs. (4.166) and (4.168), the rings of the constellation have to be discretized in phase ($\phi_k = 2\pi \frac{n}{N_\mathrm{d}}$ with $n = 0, \cdots, N_\mathrm{d}-1$), where N_d is the number of discretization steps. Thus an exhaustive maximization would involve the evaluation of $(N_\mathrm{d} N_\mathrm{r})^K$ pulses, which rapidly becomes prohibitively large. We thus restrict the tested set to some random combinations, assuming that for some number of iterations N_i the time-slot T_{N-SOL} and bandwidth-demand B_{N-SOL} is approximated sufficiently well.

We introduce the normalized TBR for multi-soliton trains to be

$$\mathrm{TBR}_{\mathrm{N-SOL,NORM}} = \frac{\mathrm{TBR}_{\mathrm{N-SOL}}}{K \cdot \mathrm{TBR}_{\mathrm{MPSK}}}, \tag{4.169}$$

where $\mathrm{TBR}_{\mathrm{N-SOL}}$ is the value computed by multiplying the numerical approximations of Eqs. (4.166) and (4.168), K is the number of discrete eigenvalues in the multi-solitons and $\mathrm{TBR}_{\mathrm{MPSK}}$ is given by Eq. (4.164). This quantity is a measure of how many resources one modulated quantity occupies compared to the case of a fundamental soliton train. Under the assumption that both systems exhibit identical MI per channel use, this would mean the overall SE would increase if $\mathrm{TBR}_{\mathrm{N-SOL,NORM}} < 1$.

Simlarly, the normalized TBP for multi-solitons can be defined as

$$\mathrm{TBP}_{\mathrm{N-SOL,NORM}} = \frac{\mathrm{TBP}_{\mathrm{N-SOL}}}{K \cdot \mathrm{TBR}_{\mathrm{MPSK}}}, \tag{4.170}$$

where now the value $\mathrm{TBP}_{\mathrm{N-SOL}}$ is a quantity obtained directly from functions $\mathcal{T}(\cdot,\cdot)$ and $\mathcal{B}(\cdot,\cdot)$ for each pulse individually. So the $\mathrm{TBR}_{\mathrm{N-SOL,NORM}}$ is the product of the

maximum values obtained over all pulse-shapes that can possibly occur using some prescribed constellation. This includes pulse-shapes that are not transmitted at the transmitter, but are generated by propagation in the fiber channel for any distance $z \in [0, \mathcal{L}]$. In the example used here, the discrete eigenvalues are purely imaginary and thus, the ring amplitudes do not change during propagation. The phases however are changing during propagation according to Eq. 4.51. This is why the maximization in Eqs. 4.166 and 4.168 is only done over all possible phases. The $\mathrm{TBR}_{\mathrm{N-SOL}}$ is thus an indicator for the required temporal transmission window and bandwidth for a certain transmitted constellation, assuming the transmission window size and receiver filter-width are fixed for all pulses in the pulse alphabet. The $\mathrm{TBP}_{\mathrm{N-SOL,NORM}}$, defined in Eq. 4.170, is the normalized TBP of individual pulses that are generated from a constellation. This values do not directly give an indication for temporal and bandwidth requirements for the whole transmission system. However, obtaining enough $\mathrm{TBP}_{\mathrm{N-SOL,NORM}}$ values for pulses that occur in the system lead to an approximation of $\mathrm{TBR}_{\mathrm{N-SOL}}$. To study the composition of pulses that occur in the system in terms of their TBP, the $\mathrm{TBP}_{\mathrm{N-SOL,NORM}}$ is more useful. In Fig. 4.24 the probability, that a multi-soliton ($K = 2$) has a certain $\mathrm{TBP}_{\mathrm{N-SOL,NORM}}$, is shown. To obtain these results,

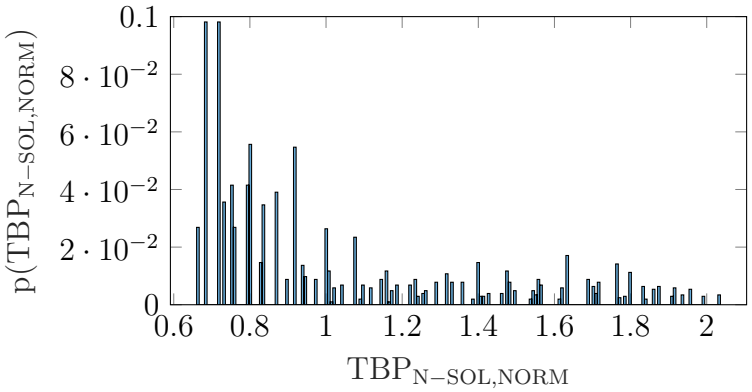


Figure 4.24.: Probability of multi-solitons having a certain TBP. ($\eta_1 = 1$, $\eta_2 = 2$, $p = 10^{-4}$, $N_{\mathrm{s}} = 256$ and $N_{\mathrm{rings}} = 4$)

each rings phase $[0, 2\pi)$ was discretized in 128 steps and the resulting constellation was used in an exhaustive TBP evaluation. As noted previously an exhaustive analysis for bigger problems quickly becomes impractical and thus an numerical evaluation of a randomly picked subset would have to be evaluated to get an approximation.

It can be seen, that for many pulses in the alphabet, the $\mathrm{TBP}_{\mathrm{N-SOL}}$ reduces in comparison to fundamental solitons. Carefully pruning the pulse-set thus could decrease

the pulse-sets TBR, potentially resulting in an increased SE. This process however would get quite involved. While the numerical results in Fig. 4.24 consider a certain set of possible phases and phase-combinations of the MPSK constellation, the generated pulses are not linked to a certain transmission distance. While generating a figure similar to Fig. 4.24 with infinitely many phases per ring would give the TBR for the respective system, it would give no indication on which pulses to use for a transmission pulse alphabet of finite size to minimize the TBR of the specific system with e.g. a fixed transmission distance $\mathcal{L}$. To facilitate this, a certain set of pulses, with as small as possible TBP for all $z \in [0, \mathcal{L}]$ has to be found. Additionally this pulse-set should ideally result in a constellation, that, as closely as possible, resembles a constellation with equidistant and equally distributed phases on each ring. It can be seen, that this can quickly results in a highly complex optimization problem.

4.6. Improved Constellations for Timing-Jitter Resistant Soliton Trains

In a communication system timing-jitter can occur due to many different effects. A prime example is imperfect clock-recovery. If some algorithm is implemented for clock recovery at the receiver, using some part of the noisy and potentially distorted received signal, a drift from the optimal timing might occur. This effect might be modeled as timing-jitter.

In the following, a strongly simplified model for timing-jitter will be used, assuming a back-to-back (B2B) transceiver scenario (i.e. the transmission channel is ideal except for the occurring timing-jitter). A received pulse $q_{\mathrm{RX}}(t)$ exhibits a slight temporal offset, compared to the transmitted pulse $q_{\mathrm{TX}}(t)$, due to timing jitter. This is modeled as

$$q_{\mathrm{RX}}(t) = q_{\mathrm{TX}}(t - \hat{t}_{\mathrm{j}}) \tag{4.171}$$

where $\hat{t}_{\mathrm{j}} \sim \mathcal{N}(0, \sigma_{\mathrm{j}}^2)$ is a Gaussian random variable with zero mean and variance σ_{j}^2. Further it is assumed, that information is to be transmitted by MPSK-modulated fundamental solitons, as described in section 4.5.1. From (4.62) it can be seen, that the modulated phases $q_{\mathrm{d}}(\lambda_1)$ are not affected by the timing-jitter, as long as the solitons discrete eigenvalue is purely imaginary and the amplitudes $|q_{\mathrm{d}}(\lambda_1)|$ are perturbed due to the occurring temporal shift. This can be checked by setting $t \to t - \hat{t}_{\mathrm{j}}$ and $\zeta_1 = 0$

in Eq. (4.62) and then absorbing the temporal shift $\hat{t}_\mathrm{j}$ into the logarithm. Another way to see this, is to evaluate the time-shift property given in Eq. (4.41) for purely imaginary discrete eigenvalues. For the following observations, it is convenient to set $|q_\mathrm{d}(\lambda_1)| = 2\eta_1 \cdot \mathrm{e}^{2\eta_1 t_\mathrm{d}}$, where t_d is the new modulated quantity. It can be seen, that now, in accordance with the observations already made in section 4.5.1, modulating the spectral amplitude by means of t_d, linearly shifts the fundamental soliton pulse in the time-domain. The resulting scheme thus is effectively a PPM transmission scheme, in which additional information is transmitted in the phase of the spectral amplitude. Since phase and amplitude modulation are decoupled under (4.171) the PPM transmission part can be studied separately.

The impairment due to timing-jitter now forms a one-dimensional AWGN channel for t_d. The capacity for the channel thus is given by Eq. (2.16) and can be stated as

$$C = \frac{1}{2}\log_2\left(1 + \frac{\mathbb{E}\left[|t_d|^2\right]}{\sigma_\mathrm{j}^2}\right). \tag{4.172}$$

If it is now assumed, that the pulse positions are modulated symmetrically around the center of the symbols time-slot the values $t_{\mathrm{d},i}$ are given by

$$t_{\mathrm{d},i} = -\frac{T_\mathrm{ppm}}{2} + \frac{iT_\mathrm{ppm}}{I-1}, \text{ for } i = 0, \cdots, I-1, \tag{4.173}$$

where $T_\mathrm{ppm} = T_\mathrm{ZC} f_\mathrm{ppm}$ is the temporal width over which the pulses position is modulated, f_ppm is defined in Eq. (4.164), T_ZC is the temporal width according to Eq. (4.64), and I is the cardinality of the pulse-alphabet used. The expected value $\mathbb{E}\left[|t_d|^2\right]$ can be calculated analytically and the resulting capacity is

$$C = \frac{1}{2}\log_2\left(1 + \frac{T_\mathrm{ppm}^2 \frac{N_\mathrm{r}+1}{12(N_\mathrm{r}-1)}}{\sigma_\mathrm{j}^2}\right), \tag{4.174}$$

with $N_r > 1$ being the number of rings in the modulation format i.e. the different positions in time the pulse can be shifted to.

In Fig. 4.25 the capacity according to (4.174) is plotted, together with some numerical results for fundamental soliton PPM over the channel described by (4.171), to verify the derived bound.

Note, that for the numerical results in Fig. 4.25, the SNR was varied by changing σ_j^2 for a fixed transmission setup. Also, for strong timing-jitter, other effects such as

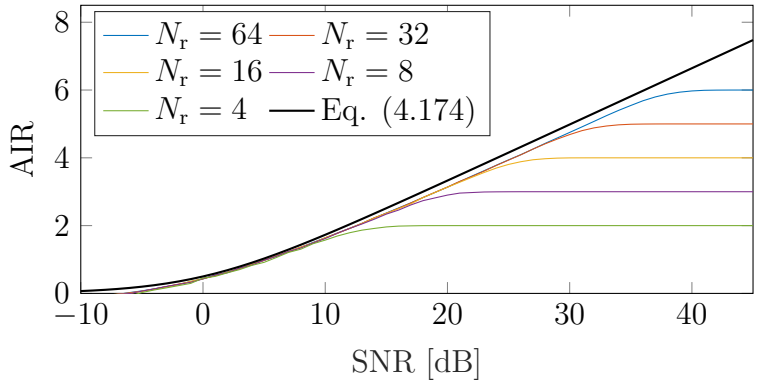


Figure 4.25.: Capacity-bound and numerical results for fundamental soliton PPM with $\eta = 1$, $p = 10^{-4}$, $f_{\text{ppm}} = 0.1$, $N_s = 1024$ and varying number of rings N_r.

ISI might further decrease the system performance. These additional effects are not captured by the simulations in Fig. 4.25.

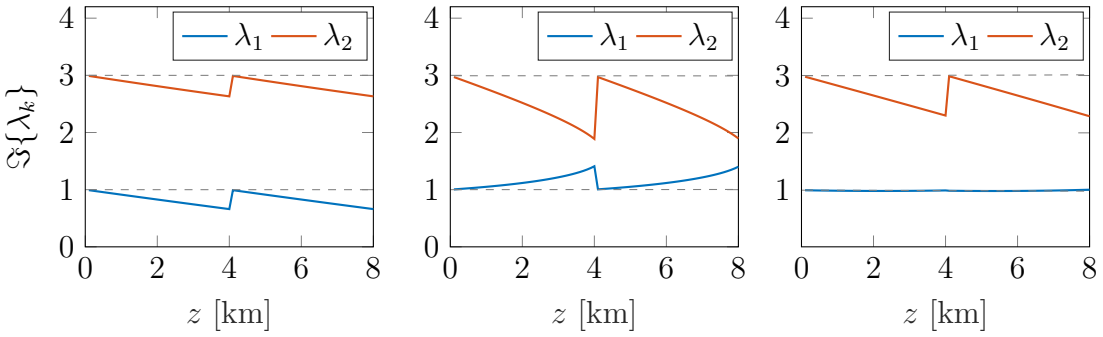
4.7. Eigenvalue Trajectories in Lossy Channels

The NFT is derived for an ideally lossless fiber model. Thus, if NFT-aided transmission is done over a channel that exhibits additional perturbations such as e.g. fiber-loss, the nonlinear quantities no longer ideally propagate according to Eqs. (4.50), (4.51) and (4.52). Also operator $\mathbf{L}(z)$ is no longer isospectral and thus e.g. discrete eigenvalues λ_k do not retain their original position.

While there are some results on the behavior of nonlinear Fourier quantities under AWGN [129], [130], there are no meaningful analytic results on the impact of fiber-loss yet. In the following, the trajectories of discrete eigenvalues in multi-solitons are studied numerically for the lossy NLSE channel. Note, that some of the span lengths used in this section are not practical and are merely used to ensure the discrete eigenvalues of the studied pulses remain trackable by the used RK-NFT. Even though, for longer span-lengths, the discrete eigenvalues generally are absorbed into the real-axis after a certain distance, this does not impose a strict limit on the amplifier spacing, since vanished eigenvalues reappear after proper re-amplification. But, beyond the point at which all discrete eigenvalues vanished, the energy of the pulse is fully shifted to the continuous spectrum and thus, the pulse exhibits dispersive behavior similar to e.g. pulses in WDM-systems.

4.7.1. Trajectory Under Lumped Amplification

The influence of propagation in a lossy NLSE channel on the two discrete eigenvalues of a multi-soliton is depicted in Fig. 4.26 for three different sets of spectral amplitudes.



(a) $q_\mathrm{d}(\lambda_2, z=0) = 1$ (b) $q_\mathrm{d}(\lambda_2, z=0) = -1$ (c) $q_\mathrm{d}(\lambda_2, z=0) \approx \mathrm{e}^{\mathrm{j}0.756\pi}$

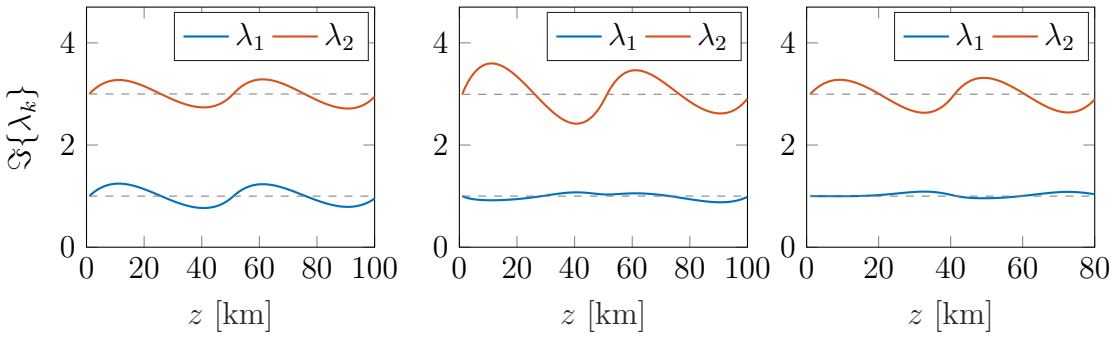
Figure 4.26.: Trajectories for $\lambda_1(z=0) = 1\mathrm{j}$, $\lambda_2(z=0) = 3\mathrm{j}$ over two EDFA amplified spans ($\mathcal{L}_\mathrm{span} = 4\mathrm{km}$) with $q_\mathrm{d}(\lambda_1, z=0) = 1$ and varying $q_\mathrm{d}(\lambda_2, z=0)$. Dashed grey lines indicate initial values for reference.

The signal is re-amplified at discrete points along the transmission link. This models amplification by e.g. EDFAs, as discussed in section 3.3.1. The pulses, for which the movement of the discrete eigenvalues is depicted in Fig. 4.26, are multi-solitons with $K = 2$. The initial value of the discrete spectral amplitude corresponding to the discrete eigenvalue with the smaller imaginary part is fixed to be $q_\mathrm{d}(\lambda_1, z=0) = 1$ for Figs. 4.26a to 4.26c, while the initial value for the second discrete spectral amplitude is varied. In Fig. 4.26a, the initial discrete spectral amplitudes are identical ($q_\mathrm{d}(\lambda_1, z=0) = q_\mathrm{d}(\lambda_2, z=0)$). It can be seen, that in this case the imaginary part of both discrete eigenvalues decreases during propagation. Since the imaginary part of the discrete eigenvalues is related to the energy in the discrete nonlinear spectrum (see Eq. (4.49)), it could also be said that, energy from both parts of the temporal pulse, corresponding to the discrete eigenvalues respectively, is transferred to the continuous nonlinear spectrum or removed from the pulse completely, due to fiber-loss. For a phase difference of π between the initial discrete spectral amplitudes (Fig. 4.26b), an interesting effect can be observed. The imaginary part of discrete eigenvalue λ_2 decreases faster than in the first case, but the imaginary part of λ_1 is increasing during propagation. This means, in addition to energy being removed form the transmission pulse due to fiber-loss, energy is additionally shifted away form the part of the pulse

corresponding to λ_2, towards the part of the pulse corresponding to λ_1. For the results shown in Fig. 4.26c, pulses with varying phase difference between the initial nonlinear spectral amplitudes $q_\mathrm{d}(\lambda_1, z=0)$, $q_\mathrm{d}(\lambda_2, z=0)$ were tested until a value was found, for which the loss of energy, due to attenuation during propagation, was compensated by the shift of energy from the λ_2 to the λ_1 portion of the pulse. In all three cases, re-amplification at $z = 4\mathrm{km}$ results in a movement of the discrete eigenvalues back into the proximity of their original location. Thus, for moderate perturbations, if the signal is properly re-amplified at regular instances, the location of the discrete eigenvalues at the receiver should be similar to their location at the transmitter. The discrete spectral amplitudes $q_\mathrm{d}(\lambda_i, z)$ however propagate according to Eq. 4.51 and thus are influenced by the exact shape of the movement exemplified in Fig. 4.26 and 4.27.

4.7.2. Trajectory Under Distributed Amplification

Similar simulations were conducted for a fiber-link with DRA amplification. In general, similar effects to the EDFA-case studied in section 4.7.1 were observed. Due to the fact, that fiber-loss is compensated in a distributed manner, the power-profile along the spans is more flat and thus it is possible to follow the trajectories of the discrete eigenvalues for longer span-lengths. In Fig. 4.27, results for three cases, similar to 4.26, are presented.



(a) $q_\mathrm{d}(\lambda_2, z=0) = 1$ (b) $q_\mathrm{d}(\lambda_2, z=0) = -1$ (c) $q_\mathrm{d}(\lambda_2, z=0) \approx \mathrm{e}^{\mathrm{j}0.8\pi}$

Figure 4.27.: Trajectories for $\lambda_1(z=0) = 1\mathrm{j}$, $\lambda_2(z=0) = 3\mathrm{j}$ over two Raman amplified fiber spans ($\mathcal{L}_\mathrm{span} = 50\mathrm{km}$) with $q_\mathrm{d}(\lambda_1, z=0) = 1$ and varying $q_\mathrm{d}(\lambda_2, z=0)$. Dashed grey lines represent initial values for reference.

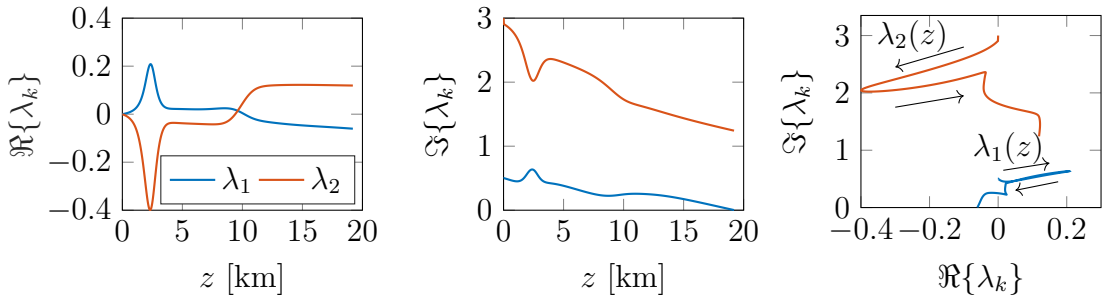
While the results depicted in Fig. 4.27a, mirror the results from Fig. 4.26a, the inverted behavior between the trajectories in Fig. 4.27b is less pronounced than in

the EDFA-case. This can be attributed to the longer span length ($\mathcal{L}_{\text{span}} = 50\text{km}$), for which the phase relation between discrete spectral amplitudes is no longer as stable as it was in the EDFA-case, where the span length was considerably lower ($\mathcal{L}_{\text{span}} = 4\text{km}$). In Fig. 4.27c, the set of initial discrete eigenvalues was again chosen such that the energy corresponding to λ_1 is held as stable as possible. Note, that once more due to the longer spans, this is no longer achieved as ideally as in Fig. 4.26c.

4.8. Collision and Absorption Phenomena

In the previous section, the span length always was chosen such, that the discrete eigenvalues remained trackable and did not get too close to each other. In the following, parameters are chosen such, that those two cases occur during propagation. Once more, there is little analytical insight so far, thus the phenomena were studied by means of numerical simulation.

In Fig. 4.28, the absorption of discrete eigenvalue λ_1 into the real axis is shown.



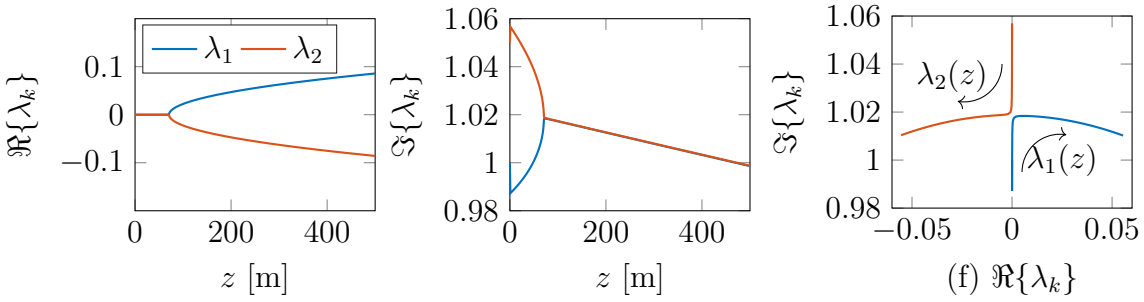
(a) Real part (b) Imaginary part (c) Complex plane

Figure 4.28.: Eigenvalue absorbtion ($q_{\text{d}}(\lambda_1) = q_{\text{d}}(\lambda_2) = 1$, $\mathcal{L}_{\text{span}} = 19.2\text{km}$, $\Delta\ell = 30\text{m}$)

To capture the data shown in Fig. 4.28, the stepsize of the SSFM was chosen to be rather small ($\Delta\ell = 30\text{m}$). The discrete eigenvalues $\lambda_1(z = 0) = 0.5\text{j}$ and $\lambda_2(z = 0) = 3\text{j}$ were chosen to speed up the absorption process of λ_1 into the real axis. Note, that for this simulation only one fiber span of length $\mathcal{L}_{\text{span}} = 19.2\text{km}$ was used without amplification. While this is not specific to the absorption phenomenon studied here, it can be seen in Fig. 4.28a, that transmission over lossy channels, also induces a change in the real part of the discrete eigenvalues. This means, that the temporal components, belonging to either λ_1 or λ_2, propagate at different velocities. In an extreme case, this could lead to the separation of the multi-soliton into its fundamental solitonic parts.

In Fig. 4.28b and 4.28c, the absorption of λ_1 into the real axis is depicted. Due to the limited capabilities of the numerical NFT-algorithms to detect discrete eigenvalues that are very close to the real axis, it is unclear what exactly happens for $\Im\{\lambda_1\} \to 0$. Thus, absorptions here means, the discrete eigenvalue is no longer trackable, due to its proximity to either the real line or spurious eigenvalues (see e.g. Fig. 4.11).

The second case, studied in this section, occurs if two discrete eigenvalues move in close proximity to each other. This "collision"-case is depicted in Fig. 4.29.



(a) Real part (b) Imaginary part (c) Complex plane

Figure 4.29.: Eigenvalue 'collision' ($q_\mathrm{d}(\lambda_1) = 1,\ q_\mathrm{d}(\lambda_2) = -1,\ \mathcal{L}_\mathrm{span} = 0.5km,\ \Delta\ell = 0.25\mathrm{m}$)

It can be seen, that the two discrete eigenvalues move towards each other during propagation, due to the choice of discrete spectral amplitudes $q_\mathrm{d}(\lambda_1) = -q_\mathrm{d}(\lambda_2)$. When their imaginary parts become similar (see Fig. 4.29b), non-zero real-parts with opposing signs are generated, as depicted in Fig. 4.29a. This effect is visible in Fig. 4.29c as well, where the "repelling" effect can be seen even clearer. Beyond the "collision" point, the two discrete eigenvalues move towards the real axis, which once again is in accordance with Eq. (4.49). Note, that the choice of discrete eigenvalues in Fig. 4.29 is rather impractical and was primarily made to force a "collision" over a short simulation distance. This is beneficial, since the step-size has to be chosen rather small to make the processes in Fig. 4.29 visible.

4.9. Path-Loss Averaged Transmission

Since fiber-loss is not considered in the channel model, for which the NFT is derived, it perturbs the continuous and discrete nonlinear spectra during propagation, as previously discussed in sections 4.7 and 4.8. To partially mitigate the negative impact

of fiber-loss on the nonlinear spectra, a NFT for the path-loss averaged (PLA) NLSE model can be used [47], [131].

4.9.1. PLA-Enhanced Transmission over EDFA Channels

For lumped EDFA amplification, as described in section 3.3.1, the propagation over one fiber-span is described by Eq. (3.1). By introducing the change of variables $Q(\tau,\ell) = \mathrm{e}^{-\frac{\alpha}{2}z}A(\tau,\ell)$ as described in [47] Eq. (3.1) can be rewritten as

$$\mathrm{j}A_z(\tau,\ell) - \frac{\beta_2}{2}A_{tt}(\tau,\ell) + \gamma\mathrm{e}^{-\alpha\ell}A(\tau,\ell)|A(\tau,\ell)|^2 = 0. \tag{4.175}$$

Under the assumption, that the envelope of $A(\tau,\ell)$ does not change too much over the length of each fiber span, the influence of the nonlinear coefficient can be averaged for the length of each span.

This average nonlinear coefficient in the lumped amplification case, can be found analytically by

$$\gamma_{\mathrm{avg}} = \frac{1}{\mathcal{L}_{\mathrm{span}}}\int_0^{\mathcal{L}_{\mathrm{span}}}\gamma\mathrm{e}^{-\alpha\ell}\mathrm{d}\ell = \gamma\frac{\mathrm{e}^{\alpha\mathcal{L}_{\mathrm{span}}}-1}{\mathrm{e}^{\alpha\mathcal{L}_{\mathrm{span}}}\alpha\mathcal{L}_{\mathrm{span}}} = \gamma\frac{G-1}{G\ln(G)}, \tag{4.176}$$

with $\ln(x)$, being the natural logarithm and G being the total fiber loss $G = \mathrm{e}^{\alpha\mathcal{L}_{\mathrm{span}}}$. With the averaged nonlinear coefficient γ_{avg} available, Eq. 4.175 can be rewritten as

$$\mathrm{j}A_z(\tau,\ell) - \frac{\beta_2}{2}A_{tt}(\tau,\ell) + \gamma_{\mathrm{avg}}A(\tau,\ell)|A(\tau,\ell)|^2 = 0. \tag{4.177}$$

If Eq. (4.177) is now normalized, as shown in section 4.1, the amplitude normalization coefficient (compare Eq. (4.3)) is now given by

$$P_0 = \sqrt{\frac{2}{\gamma_{\mathrm{avg}}L_0}}. \tag{4.178}$$

The time domain signal is thus rescaled, using γ_{avg} to mitigate the negative effects of fiber-loss during propagation.

In Fig. 4.30, the BER for transmission of a MPSK-modulated multi-soliton ($K = 2$) over an EDFA amplified link is shown with and without using the PLA method discussed above. It can be seen, that using the PLA method of rescaling transmission

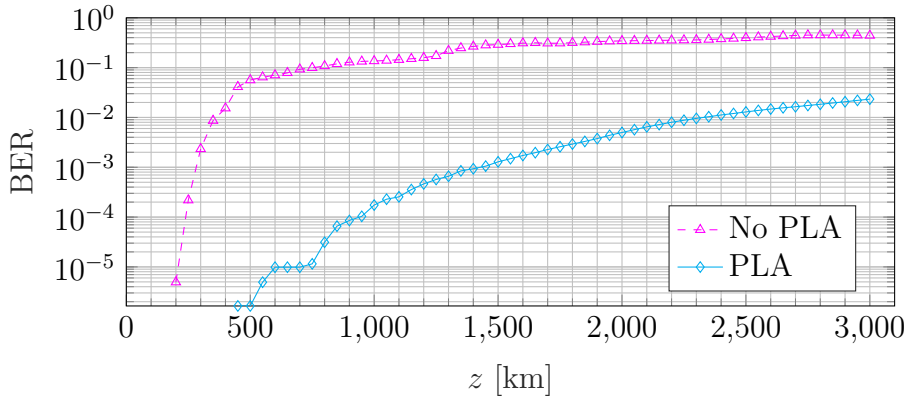


Figure 4.30.: BER with and without PLA approach (EDFA, $R_\mathrm{S} = 0.8$ GBaud, $\mathcal{L}_\mathrm{span} = 10$ km).

signals, improves the BER performance quite significantly. Here the spans were chosen to be quite short, but this improvement is present over a wide range of system settings. The underlying mechanism that results in the shown improvement, is that the pulses are linearly rescaled such that in the beginning their power is too high and thus they are exposed to more nonlinear effects than predicted by the ideal channel model for which the NFT is derived. During propagation the pulse in attenuated and thus experiences less and less nonlinear effects. On average this more closely resembles the underlying channel model, than transmitting pulses that are only optimally matched to the channel for $z = 0$. Some more results can be found in appendix D.4. The simulation parameters used to obtain the presented results are given in table C.2.

4.9.2. PLA-Enhanced Transmission over DRA Channels

The PLA approach can be used for transmission channels with DRA amplification as well [131]. The channel for the DRA-case can be modeled as

$$\mathrm{j}Q_z(\tau,\ell) - \frac{\beta_2}{2}Q_{tt}(\tau,\ell) + \gamma|Q(\tau,\ell)|^2Q(\tau,\ell) = \mathrm{j}(\underbrace{-\frac{1}{2}\alpha + \frac{1}{2}g_\mathrm{R}P_\mathrm{P}}_{g(\ell)})Q, \tag{4.179}$$

neglecting noise. It can be seen, that if fiber-loss and DRA-amplification parameters are expressed as function $g(\ell)$, the channel equation is similar to Eq. (3.1) and thus a change of variables can be done similarly to the PLA approach described for EDFAs

in section 4.9.1. Using

$$Q(\tau,\ell) = \mathrm{e}^{\int_0^\ell g(y)\mathrm{d}y} A(\tau,\ell), \tag{4.180}$$

Eq. (4.179) can be rewritten as

$$\mathrm{j}A_z(\tau,\ell) - \beta_2\frac{A_{tt}(\tau,\ell)}{2} + \gamma\underbrace{\mathrm{e}^{2\int_0^\ell g(y)\mathrm{d}y}}_{G(\ell)}A(\tau,\ell)|A(\tau,\ell)|^2 = 0. \tag{4.181}$$

Again, it is assumed that the envelope of $A(\tau,\ell)$ does not change too much over one fiber span, such that an average value γ_{avg} can be computed. Here the result depends on the exact setup of the DRA fiber-span. In the following it is assumed, that the fiber-span is pumped in co- and counter-propagating direction with equal pump-powers, as described in section 3.3.2. If now the effects of pump-depletion are negelected, the amplification parameter $G(\ell)$ is given by

$$G(\ell) = \exp\left(-\alpha_{\mathrm{S}}\ell + g_{\mathrm{R}}P_{\mathrm{P}}(\ell=0)\left(\frac{1-\mathrm{e}^{-\alpha_{\mathrm{P}}\ell}-\mathrm{e}^{-\alpha_{\mathrm{P}}\mathcal{L}_{\mathrm{span}}}+\mathrm{e}^{-\alpha_{\mathrm{P}}(\mathcal{L}_{\mathrm{span}}-\ell)}}{\alpha_{\mathrm{P}}}\right)\right). \tag{4.182}$$

The averaged nonlinearity coefficient

$$\gamma_{\mathrm{avg}} = \frac{\gamma}{\mathcal{L}_{\mathrm{span}}}\int_0^{\mathcal{L}_{\mathrm{span}}} G(\ell)\mathrm{d}\ell, \tag{4.183}$$

can be calculated numerically. This value then can be used in Eq. (4.178) to rescale the transmission pulse in the de-normalization step.

In Fig. 4.31, the BER for transmission of the same signal as in Fig. 4.30 over an DRA amplified link is shown with and without using the PLA method discussed above. The DRA scheme here uses co- and counter- propagating pump-lasers. It can be seen, that, in comparison to Fig. 4.30, the improvement using the PLA method is significantly smaller. Since the power-profile of the signal is much more flat for the DRA-case in comparison to the EDFA-case (see Fig. 3.5), it makes sense that the averaging of the nonlinear operator results in a value closer to the channels initial γ. If γ_{avg} and γ are similar, the re-scaling due to PLA will be less pronounced and thus the effect on the BER is less visible in Fig. 4.31. Some further results for the DRA channel can be found in appendix D.4. The simulation parameters used to obtain the presented results are given in table C.2.

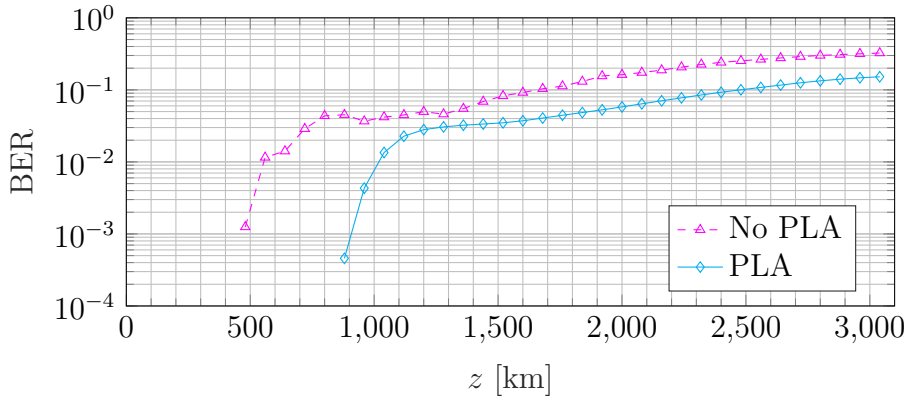


Figure 4.31.: BER with and without PLA approach (DRA, $R_\mathrm{S} = 1.2$ GBaud, $L_\mathrm{span} = 80$ km).

4.10. Fiber Parameter Estimation using the NFT

The NFT has several applications beyond its treatment in the context of optical transmission systems [89], [90]. An application that is of interest for the design of transmission systems, is the estimation of fiber channel parameters utilizing the NFT. The INFT generates time-domain signals for its nonlinear spectral components, dependent on the given channel parameters. If there is a mismatch, between the channel assumed by the NFT/INFT and the channel over which the signal is propagating, the signal is distorted. While this distortions negatively impact the data transmission capabilities of systems, they can be utilized for the estimation of one or more fiber-parameters as discussed in the following.

4.10.1. Nonlinearity Estimation with Fundamental Solitons

In this first estimation scheme, the nonlinearity parameter γ is to be estimated using fundamental solitons as probing pulses. The discrete eigenvalue λ_1 of this pulses is chosen to be purely imaginary. The discrete spectral amplitude is chosen to be $q_\mathrm{d}(\lambda_1) = 2\Im\{\lambda_1\}\mathrm{e}^{-\mathrm{j}\phi_1}$, where the magnitude $|q_\mathrm{d}(\lambda_1)|$ is chosen such that the soliton is zero-centered and the phase can be chosen arbitrarily. We set it to $\phi_1 = \frac{\pi}{2}$ to further simplify the expression for a fundamental soliton at $z = 0$ to

$$q(t, z = 0) = 2\eta\mathrm{sech}(2\eta t), \tag{4.184}$$

where $\eta = \Im\{\lambda_1\}$ is the imaginary part of the discrete eigenvalue. The corresponding pulse in natural units is given by $Q(t, z = 0) = q(t, z = 0)\sqrt{P_0}$, with P_0 either defined by Eq. (4.3) or, if the PLA scheme is used by Eq. (4.178). Since the normalization coefficient P_0 is dependent on γ or γ_{avg}, a mismatch in the nonlinearity parameter between channel and NFT/INFT leads to perturbations, since the generated pulse is no longer matched to the channel. Thus, it can be presumed that, not only the discrete eigenvalue λ_1 is no longer in its assumed position, but also some of the energy of the pulse will be dispersed during propagation. As a result, it is assumed, that the discrete eigenvalue of the pulse at the receiver deviates more strongly from the assumed position if the mismatch in the nonlinearity parameter is bigger. In the following, it is assumed that the PLA scheme is used throughout this section and thus the channel nonlinearity will simply be signified by γ_{ch}. The nonlinear parameter of the fiber-channel is assumed to be fixed and the nonlinear parameter used for the (re)scaling will be given by γ_{sc}. The factor $\delta_\gamma = \frac{\gamma_{\text{sc}}}{\gamma_{\text{ch}}}$ is introduced as a measure of the mismatch between these two variables.

In Fig. 4.32, the NMSE between the transmitted and detected discrete eigenvalues for different values of δ_γ.

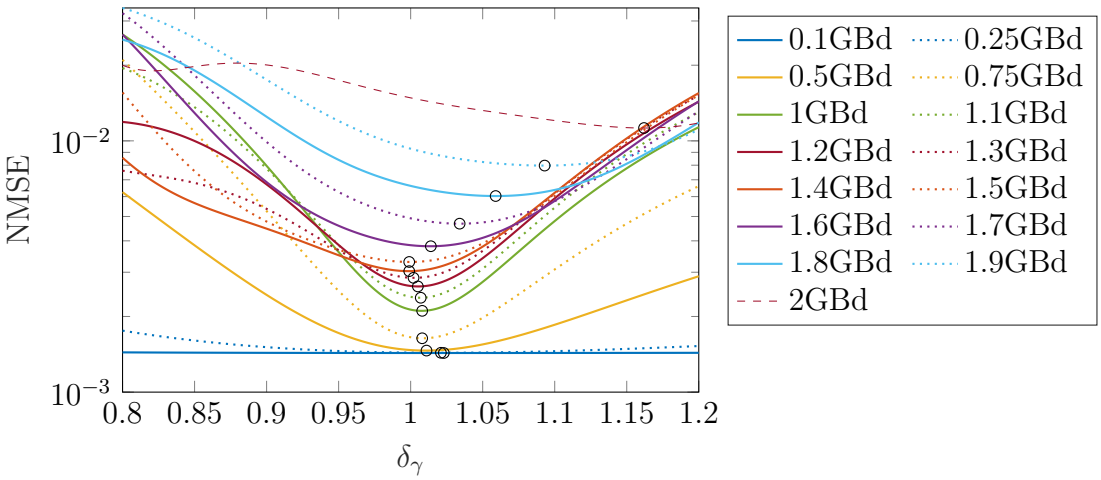


Figure 4.32.: Estimation accuracy for the fiber nonlinearity parameter γ using fundamental soliton with $\lambda_1 = 0.5\text{i}$. Noiseless Case. 20x100km

Note, that the signal power P_{S} for trains of fundamental solitons is increasing with the Baud-rate. It can be seen, that each of the obtained curves has a minimum (signified by black circles). This minimum could be seen as an estimate of the nonlinear parameter γ_{ch}, since for this value δ_γ the pulses seem to be matching the channel the

best. For a good portion of the tested Baud-rates, the minimum is quite close to $\delta_\gamma = 1$, but especially for higher Baud-rates the deviation from the assumed optimum point gets higher. Further, the slopes for very low Baud-rate curves are very shallow. This could be due to the fact, that, for very low signal powers, the amount of nonlinearity that can be measured is very low and thus the change in γ_{sc} does not influence the results significantly.

4.10.2. Joint Parameter Estimation with Multi-Solitons

Improving on the scheme described previously, the NFT-aided approach for parameter estimation can be extended to estimate both, nonlinearity and dispersion parameters of the probed fiber channel. While the exact procedure used in this thesis is slightly different, the fundamental idea for this estimation technique was presented in [132]–[134].

For easier reference, the quantities c_q and c_z are introduced. These quantities are defined as part of the normalization coefficients first introduced in Eqs. (4.2) to (4.4). They are given by

$$q(t,z) = T_0 \underbrace{\sqrt{\frac{\gamma}{|\beta_2|}}}_{c_q} Q(\tau, \ell), \tag{4.185}$$

$$z = \frac{1}{{T_0}^2} \underbrace{\frac{|\beta_2|}{2}}_{c_z} \ell. \tag{4.186}$$

The objective of the algorithm is to obtain c_q, c_z from which the fiber parameters are easily obtained.

In a first step, parameter c_q is estimated by searching for the value that results in the minimum error in the received discrete eigenvalues. Even though the probing pulses are multi-solitons, the procedure is similar to the procedure described in section 4.10.1 for γ. Note, that here it is assumed that some region in which the parameter c_q lies is known. If this is not assumed, there still is a way to define an initial search region, outlined in [133].

After an estimate on c_q is obtained, the second parameter c_z can be computed. The

normalized transmission length can be computed using Eq. (4.186). It is given by

$$\mathcal{Z} = \frac{1}{{T_0}^2} \underbrace{\frac{|\beta_2|}{2}}_{c_z} \mathcal{L}. \tag{4.187}$$

It can be seen that, since the nonlinear spectrum evolves during propagation, according to Eqs. (4.50) to (4.52), the normalized distance can be obtained by comparing the received to the transmitted values $b(\lambda_k)$ (or the discrete spectral amplitudes). It is assumed that the physical link-length is known and thus by using Eq. (4.187), the parameter c_z can be estimated.

While it is in general possible to obtain c_z just by numerically matching the transmitted and received sets, a more refined method, described in [132], will be shortly reviewed in the following. Assuming the discrete eigenvalues λ_k have a non-zero real part, the magnitudes of $b(\lambda_k)$ change according to

$$|b(\lambda_k, z = \mathcal{Z})| = |b(\lambda_k, z = 0)| \mathrm{e}^{\Re(4\mathrm{j}{\lambda_k}^2)\mathcal{Z}}. \tag{4.188}$$

Rearranging Eq. (4.188) yields

$$\mathcal{Z} = \frac{log(|b(\lambda_k, z = \mathcal{Z})|) - log(|b(\lambda_k, z = 0)|)}{\Re(4\mathrm{j}{\lambda_k}^2)}. \tag{4.189}$$

Note that, if perturbations during transmission occur the results from Eq. (4.189) for some discrete eigenvalue λ_k will just be estimates, which will be signified by $\hat{\mathcal{Z}}_k$. For the simulation results in this section, the estimates $\hat{\mathcal{Z}}_k$ were averaged by a weighted sum according to

$$\hat{\mathcal{Z}} = \frac{\sum_{k=1}^{K} w_k \hat{\mathcal{Z}}_k}{\sum_{k=1}^{K} w_k}, \tag{4.190}$$

where the weights were set according to $w_k = \Im\{\lambda_k\}$. This can be justified by the fact, that discrete eigenvalues with a higher imaginary part, correspond to a larger portion of the time-domain pulses energy [133]. Finally the fiber-parameters can be estimated, using Eqs. (4.187), (4.185) and (4.186). A flow-chart for the presented algorithm is shown in Fig. B.1 in appendix B.4.

Some simulations were conducted, to get some intuition on the accuracy of the presented method. The two estimated coefficients, introduced in this section, are thus

$c_q \approx 2.38 \cdot 10^{11}$ $\mathrm{W}^{-\frac{1}{2}}\mathrm{s}^{-1}$ and $c_z \approx 1.058 \cdot 10^{-26}$ $\mathrm{s}^2\mathrm{m}^{-1}$. The probing pulse was generated by the DT-INFT with three discrete eigenvalues $\lambda_1 = -0.1 + \mathrm{j}0.3$, $\lambda_2 = 0.1 + \mathrm{j}0.3$, $\lambda_3 = \mathrm{j}0.15$ and corresponding spectral amplitudes $q_\mathrm{d}(\lambda_1) = 1.0 + \mathrm{j}1.0$, $q_\mathrm{d}(\lambda_2) = 4.0 + \mathrm{j}4.0$, $q_\mathrm{d}(\lambda_3) = \mathrm{j}1.0$. Note that the discrete eigenvalue with zero real part is only added to the pulse to increase the signal power. This improves the measurement of the nonlinearity parameter in the first step of the parameter estimation algorithm, but can not be used to compute c_z, since Eq. (4.189) is used for its estimation. Parameter T_0 is set such, that the pulse does not move out of its allotted time-slot during propagation.

First, the algorithm was tested under ideal conditions, setting $\alpha = 0\frac{\mathrm{dB}}{\mathrm{km}}$. This lead to the results shown in Tab. 4.2. The estimation was done, using either the discrete

Table 4.2.: Estimation of fiber parameters β_2 and γ (Ideal Case).

	True value	Estimation (using $q_\mathrm{d}(\lambda_k)$)	Estimation (using $b(\lambda_k)$)
β_2 $[\frac{\mathrm{s}^2}{\mathrm{m}}]$	$-2.1173 \cdot 10^{-26}$	$-2.117785 \cdot 10^{-26}$	$-2.117765 \cdot 10^{-26}$
γ $[\frac{1}{\mathrm{W \cdot km}}]$	0.0012	0.00120029	0.00120028
Error [%]		0.02477%	0.0238%

spectral amplitudes given above, or the corresponding $b(\lambda_k)$-values. It can be seen, that the algorithm estimates the parameters very well, with slightly better performance if $b(\lambda_k)$-values are used. The measured error for both fiber parameters is identical. This is due to the fact, that the minimum error in the discrete eigenvalues is obtained at the true value c_q by the implemented grid-search. The error introduced by the estimated value c_z is thus affecting both parameters identically.

The algorithm was also tested for an EDFA-amplified fiber channel, using the PLA-method outlined in section 4.9.1. The results are given in Tab. 4.3.

Table 4.3.: Estimation of Fiber parameters β_2 and γ_avg (PLA Case).

	True Value	Estimation (using $q_\mathrm{d}(\lambda_k)$)	Estimation (using $b(\lambda_k)$)
β_2 $[\frac{\mathrm{s}^2}{\mathrm{m}}]$	$-2.1173 \cdot 10^{-26}$	$-2.2153 \cdot 10^{-26}$	$-2.1737 \cdot 10^{-26}$
γ_avg $[\frac{1}{\mathrm{W \cdot km}}]$	$3.1754 \cdot 10^{-4}$	$3.299 \cdot 10^{-4}$	$3.2389 \cdot 10^{-4}$
Error(β_2) [%]		4.629%	2.664%
Error(γ_avg) [%]		3.892%	1.999%

It can be seen, that the parameter γ_avg is smaller than the value for γ, shown in Tab. 4.2. This is due to the fact, that the channel is EDFA-amplified and thus, the signal is mostly affected by nonlinearity for the first few kilometers. After that the

signal-power has decreased significantly and nonlinear effects are almost negligible. It can be seen, that now the error for each parameter is different. This is due to the fact, that in this case c_q is no longer ideally estimated, and thus the error in γ_{avg} now is dependent on both estimated parameters c_q and c_z. The fiber parameter γ can be recovered from γ_{avg} using Eq. (4.176). Simulation parameters are given in table C.3 and further results can be found in appendix D.5.

Note, that an equation similar to Eq. (4.189) can be found, to compare angles of transmitted and received $b(\lambda_k)$ values for purely imaginary discrete eigenvalues. This would lead to a sequence of normalized length values and thus estimates for β_2. However, often a rough estimate for β_2 is known already and thus a plausible result could be picked from the sequence. The benefit of this method would be, that multi-solitons with purely imaginary discrete eigenvalues are easier to handle in a measurement setup, since they do not have components that drift out of the transmission window during propagation. It can be seen, that this parameter estimation scheme has several further interesting applications in regards to NFT-aided systems. One could be to calibrate and monitor NFDM systems. Another one could be optimizing the linear rescaling of transmission pulses beyond the capabilities of the PLA-method, by e.g. utilizing the high Baud-rate measurements in Fig 4.32 to rescale pulses. Doing so can grant a small improvement on top of the improvement gained by using the PLA-method. Some first results for this additional heuristical scaling are shown in Fig. 4.33. Since the trans-

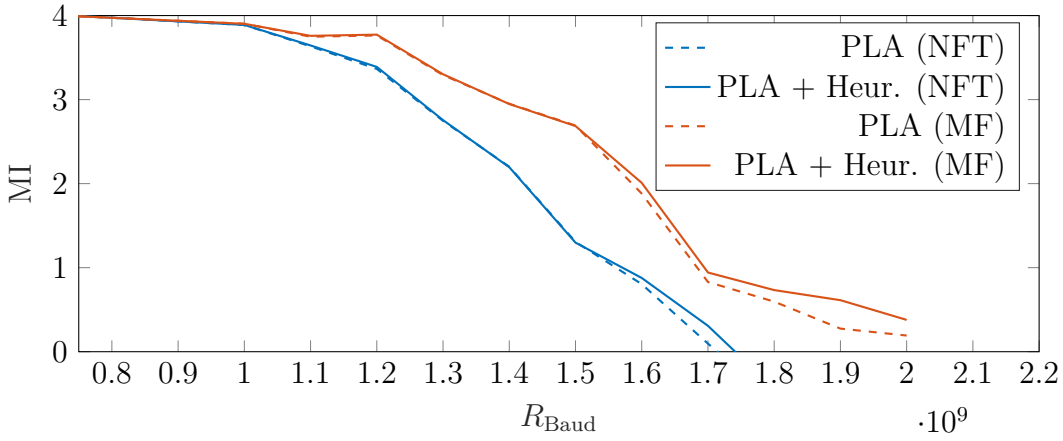


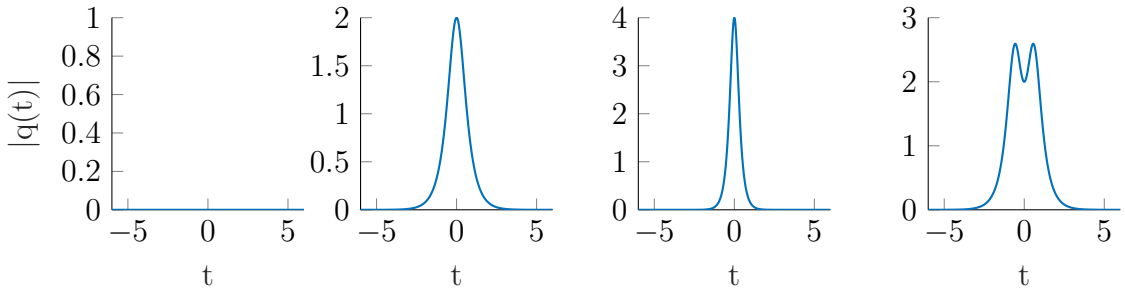
Figure 4.33.: $\mathcal{L}$ =20x100km, Multi-soliton ($K = 2$), 8-MPSK modulated $b(\lambda_k)$.

mission system utilizes multi-solitons, the signal power is linked to the symbol rate and thus higher values R_{Baud} signify higher average signal power. It can be seen, that

there is some slight improvement in the high power regime where δ_γ (compare Fig. 4.32) is deviating more strongly from its ideal value. The curves labeled NFT were measured, using an NFT algorithm at the receiver, while the curves labeled with MF are for receivers using matched-filters (MFs). The latter receiver type, which shows improved stability against noise when compared to a NFT detection scheme, will be discussed in more detail in section 4.13. While the region in which the improvement takes place here is not particularly useful, it shows that, in general, there are additional MI improvements possible using this heuristic scaling method.

4.11. Labeling Approaches for EV-OOK Systems

In eigenvalue on-off keying (EV-OOK), information is transmitted by the presence or absence of discrete eigenvalues in a multi-soliton. The discrete eigenvalues are drawn from a chosen set $\Lambda_{\mathrm{EVOOK}} = \{\lambda_1, \lambda_2, \cdots, \lambda_K\}$ of possible candidates with cardinality $|\Lambda_{\mathrm{EVOOK}}| = K$. By viewing each discrete eigenvalue as a separate channel to transmit a single bit by its presence or absence in the discrete spectrum, it can be seen, that each pulse can transmit a bit-vector $\mathbf{b}_{\mathrm{TX}} = [b_1, \cdots, b_K]$ of length K. An example of a small pulse-set is depicted in Fig. 4.34. In the following it is assumed, that all



(a) EVs: {}, $\mathbf{b}_{\mathrm{TX}} = [0, 0]$ (b) EVs: {j}, $\mathbf{b}_{\mathrm{TX}} = [1, 0]$ (c) EVs: {2j}, $\mathbf{b}_{\mathrm{TX}} = [0, 1]$ (d) EVs: {j, 2j}, $\mathbf{b}_{\mathrm{TX}} = [1, 1]$

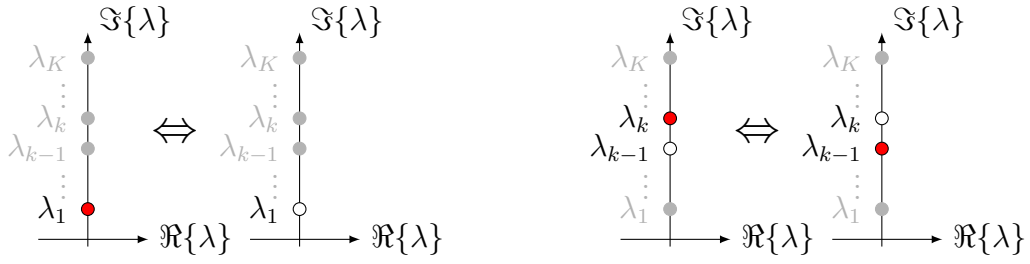
Figure 4.34.: Exemplary pulse alphabet of EV-OOK system.

discrete eigenvalues are purely imaginary and ordered such that $\Im\{\lambda_i - \lambda_{i-1}\} > 0$. The assumption that $\Re\{\lambda_k\} = 0 \,\forall\, k \in [1, K]$ is necessary for the heuristic channel model derived in the following and the condition on the imaginary part is done mostly for convenience and brevity of notation.

In the following, an abstracted and simplified heuristic channel model is studied, for the transmission of EV-OOK-modulated multi-solitons over the NLSE channel. The

model is mostly based on the observations in [135] and Eq. (4.49). The goal is to find bit-labeling schemes, that improve the BER in the aforementioned systems.

The abstract channel model, generates the error types depicted in Fig. 4.35. An



(a) Type 1 (b) Type 2

Figure 4.35.: Most probable errors in heuristic channel model (Red → can be detected. White → can not be detected.)

error of type 1, as depicted in Fig. 4.35a, occurs if the discrete eigenvalue with the smallest imaginary part vanishes into the real line. This effect was already studied in section 4.8. The second error type, depicted in Fig 4.35b, occurs if a discrete eigenvalue is detected as another one which was not originally sent. In abstract terms, it could be described as a neighboring 'absent' and 'present' state switching places. The intuition behind the model stems from the idea, that, for moderate power-fluctuation, the discrete eigenvalues will most likely be detected falsely as some of their direct neighbors (compare Figs. 4.26 and 4.28b) or the smallest one will vanish into the real line (compare Fig. 4.28).

Note, that energy can also be transferred to the continuous spectrum without changing the overall energy of the time domain pulse. From previous observations in sections 4.7 and 4.8, it seems that this process does not occur over much smaller distances z, than the exchange in the discrete spectrum and the general energy reduction due to fiber-loss. Thus, the exchange of energy between continuous and discrete spectrum was not included in the tested model. For conventional modulation schemes such as QAM in a WDM-system, gray labeling [136] is used to minimize the number of occurring bit errors per symbol error. Assuming that, for moderate perturbations, only errors between nearest neighbors occur only one bit error per symbol error occurs if gray labeling is used. Such a labeling method, minimizing the bit-errors per symbol-error, is also desirable for EV-OOK modulation.

The naive approach to labeling, is to use the binary vectors $\mathbf{b}_{\mathrm{TX}}$ as a direct indicator

for the presence or absence of a discrete eigenvalue in the pulse. This means that $b_k = 1$ if $a(\lambda_k) = 0$ and $b_k = 0$ if $a(\lambda_k) \neq 0$. This scheme will be called *binary* labeling in the following. It can be seen directly, that if an error of type 2 occurs and thus discrete eigenvalue λ_k is detected as $\lambda_{k\pm1}$, two adjacent bit errors occur in positions k and $k \pm 1$.

4.11.1. 'Gray' Labeling

The second labeling scheme will be called 'gray'-labeling in the following. While it is not similar to the well known gray-labeling for e.g. QAM symbols from an algorithmic standpoint, it achieves the same goal, which is to cause errors between adjacent symbols to only generate one bit-error. Here adjacent means the one symbol turns into another by either one error of type 1 or one error of type 2. The labels can be assigned according to

$$\mathbf{u}_{\mathrm{TX}} \cdot \mathbf{G}_{\mathrm{gray}} = \mathbf{b}_{\mathrm{TX}}, \tag{4.191}$$

with

$$\mathbf{G}_{\mathrm{gray}} = [\mathbf{d}_K,\ \mathbf{d}_{K-1},\ \cdots,\ \mathbf{d}_1], \tag{4.192}$$

where

$$\mathbf{d}_k = \begin{cases} [1, \mathbf{0}_{K-1}]^{\mathrm{T}}, & \text{for } k = K \\ [\mathbf{0}_{K-k-1}, 1, 1, \mathbf{0}_{k-1}]^{\mathrm{T}}, & \text{otherwise} \end{cases}, \tag{4.193}$$

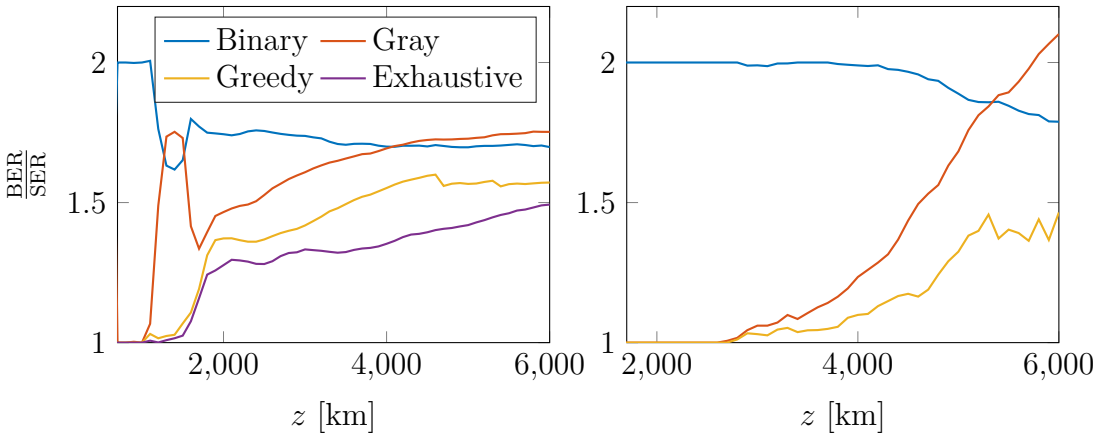
and $\mathbf{u}_{\mathrm{TX}}$ is the gray bit-label. In the above equations $\mathbf{0}_k$ and $\mathbf{1}_k$ are zero and one vectors of length k. The proof can be found in appendix A.4.

4.11.2. Greedy Labeling

Another labeling scheme, that was implemented and tested for EV-OOK-systems, is a simple form of greedy labeling. This algorithm labels the symbols according to a measured transition probability matrix and is thus relying on the use of pilot-data. Symbol-pairs are labeled such, that the ones with the highest transition probability have bit-labels that only differ in one bit. As already described in detail in section 3.4.3, this can lead to suboptimal labeling of symbol-pairs later in the process. The exact algorithm used is given in pseudo-code in algorithm 2 in appendix B.2.

4.11.3. Simulation Results

The labeling schemes presented beforehand, were evaluated numerically for transmission over a DRA-amplified fiber-channel. The span-length was set to be $\mathcal{L}_{\text{span}} = 100\,\text{km}$ and the total transmission distance was $\mathcal{L} = 6000\,\text{km}$. In Fig. 4.36, the bit-error per symbol-error for two different scenarios is given. For the additional curve for $K = 3$



(a) $K = 3$, $R_{\text{Baud}} = 1.2\,\text{GBaud}$ (b) $K = 5$, $R_{\text{Baud}} = 0.57\,\text{GBaud}$

Figure 4.36.: Bit-error per symbol-error for several bit-labeling schemes

labeled 'Exhaustive', all possible label combinations were tested and the one with the lowest bit-error per symbol-error was used. It gives a lower bound on the bit-error per symbol-error achievable for this specific realization. The complexity of this method is very high, thus it could no longer be used for $K = 5$.

For Fig. 4.36a it can be seen that, for the most part, the binary labeling scheme is outperformed by the other algorithms. Greedy labeling also consistently outperforms the 'gray'-labeling approach, although it should be mentioned that the drawback of the greedy labeling method is, that pilot-data has to be used to compute the transition probability matrix. The cause for the unexpected behavior of the binary and gray curve around $1000\,\text{km} - 1500\,\text{km}$ is not fully understood. The effect could be caused by some unfavorable breathing-motion of the pulses in this region of the fiber. However, this effect is not visible in the greedy and exhaustive curves. Since they are adaptive algorithms that do not rely on the assumptions made in the heuristic channel model (Fig. 4.35) for the derivation of the gray labeling algorithm, they seem to not show the same degradation of performance if the channel behavior deviates from this model. Most interestingly, for low BER values which correspond to small values of z (compare

results in appendix D.6), all algorithms exhibit a bit-error to symbol-error ratio of $\frac{\mathrm{BER}}{\mathrm{SER}} = 1$, except for the binary labeling scheme, which exhibits a value of $\frac{\mathrm{BER}}{\mathrm{SER}} = 2$. The reason for the binary scheme having this value for small perturbations was already given above. For the gray algorithm, this indicates that the heuristic channel model in Fig. 4.35 seems to capture the errors occurring in the transmission system well for small perturbations. Thus, the gray labeling scheme seems to work as intended in this region.

For $K = 5$ (see Fig. 4.36b), the results give similar insight, lacking the irregular behavior visible for $K = 3$. It can be seen, that the optimal value of $\frac{\mathrm{BER}}{\mathrm{SER}} = 1$ is achieved by gray and greedy labeling for a distance of up to 3000 km. Note, that this increase in reach for the improvement of bit-error to symbol-error ratio probably has to be attributed to the lower Baud-rate used for this simulation. In addition to the plots shown in this section, some more results can be found in appendix D.6.

4.12. Clustering Algorithm for Discrete Spectrum Data

Consider an NFT communication system that transmits N_{sym} symbols of information on the values $b_k = b(\lambda_k)$ of a solitonic pulse with K eigenvalues. Assuming that each b_k is taken from a constellation of N_{c} points, there are $M = N_{\mathrm{c}}^K$ possible transmit symbols, or constellation points $\boldsymbol{b}_m = (\Re\{b_1[m]\}, \cdots, \Re\{b_K[m]\}, \Im\{b_1[m]\}, ..., \Im\{b_K[m]\})^{\mathrm{T}}$ for $m \in \{1, \cdots, M\}$. At each time index $n \in 1, ..., N_{\mathrm{sym}}$, one constellation point $\boldsymbol{b}_m$ is chosen, and the symbol

$$\boldsymbol{x}_n = \boldsymbol{b}_m = (\Re\{x_1[n]\}, \cdots, \Re\{x_K[n]\}, \Im\{x_1[n]\}, ..., \Im\{x_K[n]\})^{\mathrm{T}}, \tag{4.194}$$

is transmitted.

Let $\boldsymbol{y}_n$ be the received symbol at time index n. It can be modeled as a RV $\boldsymbol{y}$ whose probability density function (PDF) is a mixture of Gaussians (MOG)

$$p_{\boldsymbol{y}}(\boldsymbol{y}) = \sum_{m=1}^{M} P_m(m) N(\boldsymbol{y}|\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) \tag{4.195}$$

where $P_m(m)$ is the probability that the m-th constellation point b_m has been sent, equal to $\frac{1}{M}$ for uniformly distributed constellation points. The vector $\boldsymbol{\mu}_m = \mathbb{E}[\boldsymbol{y}|\boldsymbol{x} = \boldsymbol{b}_m]$ is the expectation of the received symbol given that the constellation point $\boldsymbol{b}_m$ was

sent, and $\boldsymbol{\Sigma}_m$ is the corresponding covariance matrix. Further, $N(\boldsymbol{y}|\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)$ is the multivariate Gaussian PDF with mean $\boldsymbol{\mu}_m$ and covariance matrix $\boldsymbol{\Sigma}_m$.

If a sufficient amount of received constellation points is available, the parameters of the MOG model can be estimated by the expectation maximization (EM) algorithm. EM [137], [138] is an iterative algorithm to find a set of parameters $\Xi_{\text{opt}} = (\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)$ (which contains all the $\boldsymbol{\mu}_m$ and $\boldsymbol{\Sigma}_m$ for $m = 1, \cdots, M$) that maximizes the probability of the received signal $\boldsymbol{Y} = (\boldsymbol{y}_1, ..., \boldsymbol{y}_{N_{\text{sym}}})$. Thus, the estimation problem is

$$\Xi_{\text{opt}} = \underset{\Xi}{\text{argmax}}(p(\boldsymbol{Y}|\Xi)) \tag{4.196}$$

with Ξ being the set of parameters over which the maximization is done. The l-th iteration of EM is comprised of an expectation step and a maximization step.

Expectation Step

For each received constellation point $\boldsymbol{y}_n$, the expectation step computes the a posteriori probabilities

$$\gamma_{nm}^{(\ell)} = p(\boldsymbol{x}_n = \boldsymbol{b}_m|\boldsymbol{y}_n) \tag{4.197}$$

$$= \frac{P_m^{(l)}(m) N\left(\boldsymbol{y}_m|\boldsymbol{\mu}_m^{(l-1)}, \boldsymbol{\Sigma}_m^{(l-1)}\right)}{\sum_{m'=1}^{M} P_m^{(l-1)}(m') N\left(\boldsymbol{y}_n|\boldsymbol{\mu}_{m'}^{(l-1)}, \boldsymbol{\Sigma}_{m'}^{(l-1)}\right)} \tag{4.198}$$

which are often called *responsibilities* [138]. Notation $a^{(l)}$ denotes the value of a at the l-th iteration.

Maximization Step

In the subsequent maximization step, the parameters $\Xi^{(l)}$ are updated for all m according to

$$N_m^{(l)} = \sum_{n=1}^{N} \gamma_{nm}^{(l)} \tag{4.199}$$

$$P_m^{(l)}(m) = \frac{N_m^{(l)}}{N_{\text{sym}}} \tag{4.200}$$

$$\boldsymbol{\mu}_m^{(l)} = \frac{1}{N_m^{(l)}} \sum_{n=1}^{N_{\text{sym}}} \gamma_{nm}^{(l)} \boldsymbol{y}_n \tag{4.201}$$

$$\boldsymbol{\Sigma}_m^{(l)} = \frac{1}{N_m^{(l)}} \sum_{n=1}^{N_{\mathrm{sym}}} \gamma_{nm}^{(l)} (\boldsymbol{y}_n - \boldsymbol{\mu}_m^{(l)})(\boldsymbol{y}_n - \boldsymbol{\mu}_m^{(l)})^{\mathrm{T}}. \tag{4.202}$$

The parameters computed in the maximization step are then fed back to the next expectation step during iteration $l+1$. Note that the algorithm also estimates the input distribution $P_m(m)$, and is therefore also useful for non-uniform input distributions. The algorithm terminates if either an appropriate difference measure between subsequently computed parameter sets is smaller than some threshold value or after a maximum number of iterations has been reached. The EM algorithm is guaranteed to converge to a (local) maximum [138]. The EM algorithm classifies the received symbols into M clusters, which correspond to the possible constellation points. The task of assigning the clusters to the constellation points is an instance of the assignment problem, which was solved using the optimal Hungarian method [139]. This method finds an assignment $f : \{1, \cdots, M\} \to \{1, \cdots, M\}$ that minimizes the sum of Euclidean distances between the cluster means $\boldsymbol{\mu}_m$ and their corresponding constellation point $b_{f(m)}$. A short example of the iterative steps of the EM algorithm is given in appendix D.7.1.

The EM algorithm was tested, using received data from a multi-soliton transmission system with $K = 2$ and a bi-directional pump Raman-amplified channel of length $\mathcal{L} = 1800\,\mathrm{km}$, which comprises $N_{\mathrm{span}} = 60$ spans of length $\mathcal{L}_{\mathrm{span}} = 25\,\mathrm{km}$. The discrete eigenvalues were chosen to be $\lambda_1 = 1.5\mathrm{j}$ and $\lambda_2 = 1\mathrm{j}$. Information was transmitted by MPSK-modulated $b_k = b(\lambda_k)$ values with $N_{\mathrm{r}} = 4$ rings and $N_{\mathrm{p}} = 4$ uniformly spaced phases per ring $(0, \frac{\pi}{2}, \pi, \frac{3\pi}{2})$, using gray-mapping for bit-labeling. The ring amplitudes were heuristically chosen to be the values $|b_1| \in \{0.0163, 0.2534, 3.9462, 61.4517\}$ and $|b_2| \in \{0.0642, 0.4004, 2.4972, 15.5724\}$ similar to [140]. Further parameters are given in Tab. C.4 in appendix C.4.

To account for the effects of ISI, the propagation was simulated in blocks of 64 modulated pulses. At the receiver, after band-pass filtering a FB TD-NFT was performed to compute the received spectral amplitudes $y_k[n]$. The $y_k[n]$ were mapped to the constellation points using (a) minimum Euclidean distance (MD) or (b) the EM algorithm. For the EM algorithm, different training set sizes N_{tr} between 6912 and 25600 symbols were used (note that the method is blind and the training set is used to transmit data). After training, the EM parameters Ξ were kept fixed and further tested on new symbols. The total number of symbols (training plus testing) was 90112 in all cases. The BER results are plotted in Fig. 4.37 as a function of link length $\mathcal{L}$.

As expected, the performance of EM depends on the size of the training set, N_{tr}. It

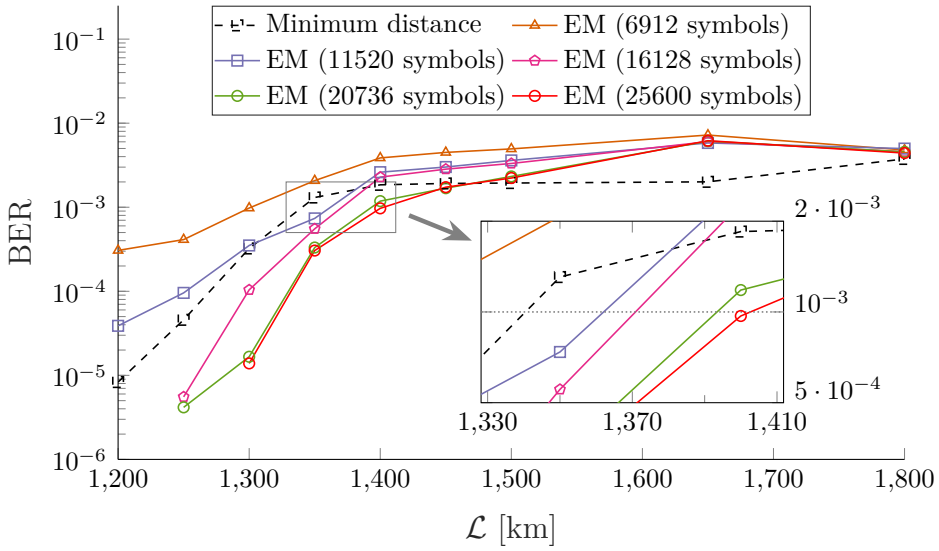
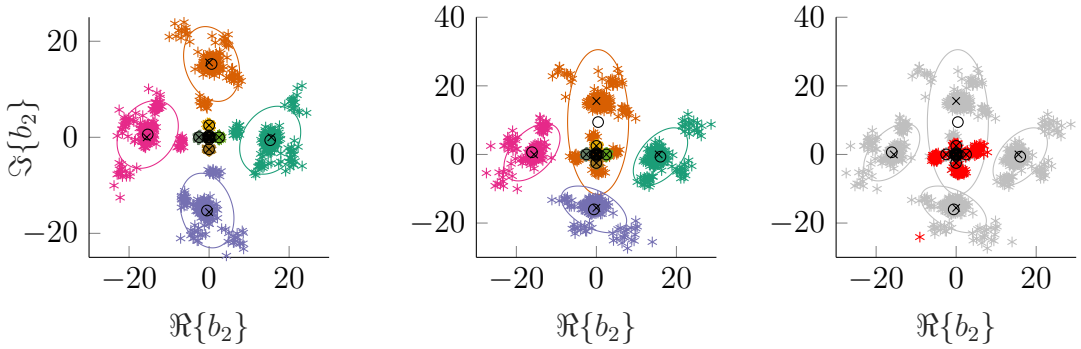


Figure 4.37.: Bit error rate for minimum distance (dashed) and EM (solid) detection for different training set sizes.

can be observed, that different symbols become unreliable at different distances. This often results in staircase-like BER curves, which could explain the seemingly saturating BER curves in Fig. 4.37. At BER $= 10^{-3}$, with a training set size of $N_{\mathrm{tr}} = 16128$ symbols, the reach of EM is 30 km larger than that of MD. At about $N_{\mathrm{tr}} = 20000$ training symbols, the performance of EM saturates at 60 km larger reach than MD.

Fig. 4.38 shows the obtained clusters for two different transmission distances and $N_{\mathrm{tr}} = 25600$. The originally transmitted positions are shown by black crosses, the black circles show the average values of each cluster and the circles are the respective 95% confidence intervals. In Fig. 4.38a the received symbols for a transmission distance of $\mathcal{L} = 1500$ km are depicted, colored by cluster. It can be seen, that even though the received values are perturbed strongly, the EM algorithm is able to form plausible clusters. In Fig. 4.38b for $\mathcal{L} = 1650$ km it can be seen, that the EM algorithm generates one cluster (orange), that results in a lot of errors (red marks in Fig. 4.38c) around the center of the plot.



(a) Clusters for b_2 ($\mathcal{L} = 1500\,\mathrm{km}$)

(b) Clusters for b_2 ($\mathcal{L} = 1650\,\mathrm{km}$)

(c) Errors for b_2 ($\mathcal{L} = 1650\,\mathrm{km}$)

Figure 4.38.: EM-clusters of received symbols.

4.13. Soliton Receivers with Matched-Filters

In all previous scenarios, the transmission pulses in the respective NFDM-systems were detected using some NFT algorithm at the receiver. If the set of possible received pulse-shapes (without perturbations) is known at the receiver, the information can be retrieved by using a bank of MFs as well. For this to be feasible, the transmission distance $\mathcal{L}$ has to be known at the receiver, since pulses change shape during propagation. Also the size of the pulse-alphabet at the receiver should not be too big for this approach, to keep the detection schemes complexity manageable. To evaluate the impact of MF-based detection on the transmission schemes, fundamental soliton pulses were used. Data was transmitted by a combination of discrete eigenvalue modulation $\lambda_1 = \mathrm{j}k$ for $k = \{1, \cdots, K\}$ and PSK-modulated spectral amplitudes.

First, the MI for transmission over an AWGN-channel was determined by numerical simulation. The results are depicted in Fig. 4.39. It can be seen, that the MI curves of the MF detector are approaching the theoretical bound for low SNR and the MI for NFT detection drops faster when the SNR decreases. The MF curves approaching the theoretical bound for small SNR is somewhat expected, since the system, except for the pulse generation method, is analog to systems for which the given bound is derived. The MF-curves deviate from the bound for higher SNR. This could be explained by the choice of transmission pulses, which were generated according to some chosen nonlinear spectral data (compare Fig. 3.9 and Eq. 4.62). Note, that for the NFT approach no previous matched filtering can be done, which increases the influence of noise on the signal and additionally the MI might be decreased by errors due to the

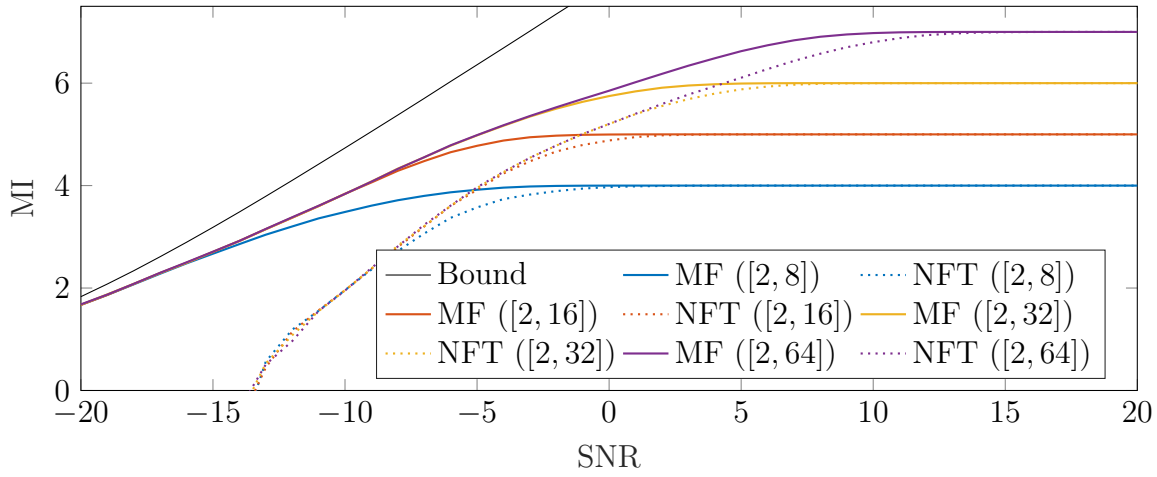


Figure 4.39.: MI for fundamental soliton discrete eigenvalue modulation + PSK modulated spectral amplitudes and detection by MF or NFT. (Legend: $[K, N_{\text{phases}}]$)

numerical behavior of the NFT algorithm. Also it was shown that AWGN noise in the linear domain does not lead to white noise in the NFD, which also impacts the MI.

Another simulation, modeling transmission over an EDFA channel where noise, loss and non-linearity influence the signal during propagation, was conducted. The results

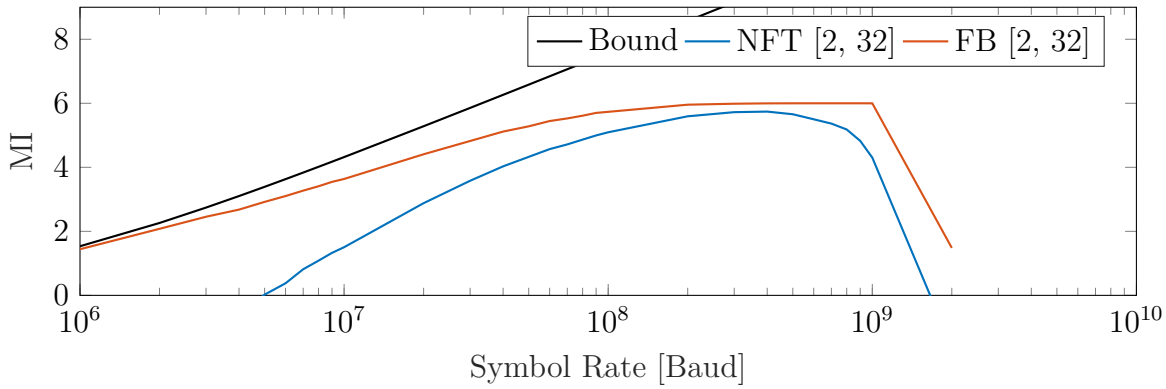


Figure 4.40.: MI for fundamental soliton discrete eigenvalue modulation + PSK modulated spectral amplitudes and detection by MF or NFT.

given in Fig. 4.40 are in this case plotted over the symbol rate, since signal power is directly linked to symbol rate for fundamental solitons. It can be seen, that in this case the MF approach significantly improves upon the NFT method as well. Again an upper bound is given. This bound was derived, for an AWGN channel (similar to

bound in Fig. 4.39), with noise power identical to the accumulated EDFA noise in the channel.

5

The Nonlinear Fourier Transform for Multi-Mode Fibers

The most prominently used channel model for the study of NFT-aided optical communication systems is Eq. (3.1), but the NFT is generally applicable in the solution of a wide variety of equations [90]. In regards to optical communication systems, several recent publications were studying the applicability of the NFT to the dual-polarization NLSE case [141]–[143]. In this chapter, this theory is expanded to the general strong coupling multi-mode case described by the SCME (3.30).

Similar to the treatment in chapter 4, the channel model is normalized first. The normalized SCME is obtained as

$$\mathrm{j}\boldsymbol{q}_z(t,z) = \boldsymbol{q}_{tt}(t,z) + \mathrm{j}2\boldsymbol{q}(t,z)\sum_{m=1}^{M}|q_m(t,z)|^2, \tag{5.1}$$

where $\boldsymbol{q}_z$ is the vector holding z-derivatives for the signals in each mode, $\boldsymbol{q}_{tt}$ is holding second order derivatives in normalized time t, $q_m(t,z)$ is the m-th entry of $\boldsymbol{q}(t,z)$ and M is the number of excited modes. The normalization is done using relations $\boldsymbol{q}(t,z) = \boldsymbol{Q}(\tau,\ell)/P_0$, $t = \tau/T_0$ and $z = \ell/L_0$, with normalization parameter relations $P_0 = \sqrt{2/(\gamma\kappa L_0)}$ and $L_0 = 2{T_0}^2/|\bar{\beta}_2|$. $\boldsymbol{Q}(\tau,\ell)$ is the vector holding the signal in natural units (compare Eq. (3.30)). It can be seen, that once more one of the normalization parameters can be chosen freely.

5.1. Basic Concept and Transformations

Starting with the eigenvalue equation of the Lax-operator $\mathbf{L}(z)$ according to Eq. (4.5) and the propagation equation for eigenvectors in z (see Eq. (4.6)), the basis for the NFT for the multi-mode case can be derived in close relation to the steps taken in section 4.1.

In a first step, the Lax pair $\mathbf{L}(z)$, $\mathbf{M}(z)$ for Eq. (5.1) has to be found. Starting from the Lax-operators given in [144], [90, Ch.2] for the dual-polarization case and using the extensions in [145] the $\mathbf{L}(z)$-operator is found to be

$$\mathbf{L}(z) = \mathrm{j}\begin{bmatrix} D & -q_1(t,z) & -q_2(t,z) & \cdots & -q_M(t,z) \\ -q_1^*(t,z) & -D & 0 & \cdots & 0 \\ -q_2^*(t,z) & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -q_M^*(t,z) & 0 & \cdots & 0 & -D \end{bmatrix}. \tag{5.2}$$

Using the framework in [146, Sec. 4] the $\mathbf{M}(z)$-operator becomes a $(M+1)$x$(M+1)$ square-matrix with entries

$$M_{11} = \mathrm{j}2\lambda^2 - \mathrm{j}\sum_{m=1}^{M}|q_m(t,z)|^2 \tag{5.3}$$

$$M_{kk} = -\mathrm{j}2\lambda^2 - \mathrm{j}|q_{k-1}(t,z)|^2 \qquad \text{for } k = 2,\cdots,M+1 \tag{5.4}$$

$$M_{1k} = -2\lambda q_{k-1}(t,z) - \mathrm{j}q_{k-1,t}^*(t,z) \qquad \text{for } k = 2,\cdots,M+1 \tag{5.5}$$

$$M_{k1} = 2\lambda q_{k-1}(t,z) - \mathrm{j}q_{k-1,t}(t,z) \qquad \text{for } k = 2,\cdots,M+1 \tag{5.6}$$

$$M_{kl} = \mathrm{j}q_k(t,z)q_l^*(t,z) \qquad \text{for } k,l = 2,\cdots,M+1,\ k \neq l. \tag{5.7}$$

While, a treatment of the NFT with periodic boundary conditions might exist for the multi-mode case, the signals are once more, supposed to have limited energy and vanish sufficiently fast for $t \to \pm\infty$ over the course of this thesis. The eigenvalue equation for operator $\mathbf{L}(z)$ can then be solved at $\pm\infty$ and boundary conditions can

be set to

$$\boldsymbol{\psi}_{\pm}^{(0)}(t,\lambda,z) \to \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \mathrm{e}^{-\mathrm{j}\lambda t}, \ \boldsymbol{\psi}_{\pm}^{(m)}(t,\lambda,z) \to \begin{bmatrix} 0 \\ \delta_1(m) \\ \vdots \\ \delta_M(m) \end{bmatrix} \mathrm{e}^{\mathrm{j}\lambda t}, \tag{5.8}$$

with

$$\delta_l(m) = \begin{cases} 1 & \text{for } m = l \\ 0 & \text{for } m \neq l \end{cases} \quad \text{for } l, m \in 1, \cdots, M. \tag{5.9}$$

It can be seen, that Eq. (5.8) is the expansion of the conditions for the two-polarization case given in [141]. They can be be used to define nonlinear Fourier coefficients via the projection equation

$$\boldsymbol{\psi}_{-}^{(0)}(t,\lambda,z) = a(\lambda)\boldsymbol{\psi}_{+}^{(0)}(t,\lambda,z) + \sum_{m=1}^{M} b_m(\lambda)\boldsymbol{\psi}_{+}^{(m)}(t,\lambda,z). \tag{5.10}$$

The nonlinear Fourier coefficients, for some fixed value z, then consist of parameter $a(\lambda)$ and M parameters $b_m(\lambda)$. Note that each parameter $b_m(\lambda)$ in general is related to all time-domain signals $q_m(t,z)$, since they are coupled during propagation by fiber-nonlinearity. Parameters $a(\lambda)$, $b_m(\lambda)$ for $m = 1, \cdots, M$ are again time independent. This facilitates the computation of the nonlinear Fourier coefficients via numerical integration of the linear propagation equation for eigenvectors in t.

The nonlinear Fourier transform for some fixed distance parameter z can be stated as

$$\mathrm{NFT}(\boldsymbol{q}(t)) = \begin{cases} q_{\mathrm{c}}^{(m)}(\lambda) = \frac{b_m(\lambda)}{a(\lambda)} & \lambda \in \mathbb{R} \\ q_{\mathrm{d}}^{(m)}(\lambda_k) = \frac{b_m(\lambda_k)}{a_\lambda(\lambda_k)} & \lambda_k \in \mathbb{C}^+ | a(\lambda_k) = 0 \end{cases}, \tag{5.11}$$

where $q_{\mathrm{c}}^{(m)}(\lambda)$, $q_{\mathrm{d}}^{(m)}(\lambda_k)$ denote the continuous and discrete nonlinear spectral amplitudes and λ_k for $k = 1, \cdots, K$ are the discrete eigenvalues in the upper half of the complex plane. Entries $q_{\mathrm{c}}^{(m)}(\lambda)$ and $q_{\mathrm{d}}^{(m)}(\lambda_k)$ form the vector $\boldsymbol{q}_{\mathbf{c}}(\lambda)$, $\boldsymbol{q}_{\mathbf{d}}(\lambda_k)$ respectively. Note that neither the operators $\mathbf{L}(z)$, $\mathbf{M}(z)$, nor the boundary conditions are defined unambiguously. While some of the possible representations might be favorable over others due to e.g. the numerical stability of the resulting algorithms, in general any definition that is valid in itself is sufficient as long as all the used transformations use the same set of assumptions.

5.2. Properties of the Nonlinear Spectra

In the following, some of the basic properties from [43] are studied and (if possible) extended to the multi-mode case. Distance parameter z is considered to be fixed and thus omitted in all sections except for the treatment in section 5.2.5. Using $y_m(t,\lambda) = \frac{v_{m+1}(t,\lambda)}{v_1(t,\lambda)} \cdot \mathrm{e}^{-2\mathrm{j}\lambda t}$, where $v_m(t,\lambda)$ are the entries of $\boldsymbol{v}(t,\lambda) = \boldsymbol{\psi}_-^{(0)}(t,\lambda)$ the continuous forward transformation can be restated as

$$\frac{\mathrm{d}y_j(t,\lambda)}{\mathrm{d}t} + y_j(t,\lambda) \sum_{m=1}^{M} q_m(t) y_m(t,\lambda) \mathrm{e}^{2\mathrm{j}\lambda t} + q_j^*(t) \mathrm{e}^{-2\mathrm{j}\lambda t} = 0 \tag{5.12}$$

$$y_j(t \to -\infty, \lambda) = 0. \tag{5.13}$$

Similarly from $\lim_{t\to\infty} f(t,\lambda) = a(\lambda)$ one can find the second order differential equation

$$\frac{\mathrm{d}^2 f(t,\lambda)}{\mathrm{d}t^2} - \left(2\mathrm{j}\lambda + \sum_{m=1}^{M} \frac{q_k'(t)}{q_k(t)}\right) \frac{\mathrm{d}f(t,\lambda)}{\mathrm{d}t} + \tag{5.14}$$

$$\sum_{m=1}^{M} |q_m(t)|^2 f(t,\lambda) + \sum_{m=1}^{M} \sum_{j=1, j\neq m}^{M} \frac{q_m'(t)}{q_m(t)} q_j(t) u_j(t,\lambda) = 0 \tag{5.15}$$

where

$$\frac{\mathrm{d}u_j(t,\lambda)}{\mathrm{d}t} - \mathrm{j}\lambda + q_j^*(t) f(t,\lambda) \mathrm{e}^{-\lambda t} = 0 \tag{5.16}$$

$$u_j(t \to -\infty, \lambda) = 0 \tag{5.17}$$

$$f(t \to -\infty, \lambda) = 1 \tag{5.18}$$

$$\lim_{t\to-\infty} \frac{\mathrm{d}f(t,\lambda)}{\mathrm{d}t} = 0. \tag{5.19}$$

The derivations for the differential representations above and all statements made in the following subsections can be found in appendix A.5.

5.2.1. Constant Phase Change

The property of a constant phase change in time domain resulting in an inverse phase change in the nonlinear Fourier domain still holds for the multi-mode case. It even can be extended to the case, where the time domain signal in each mode is shifted by

a different constant phase. This can be stated as

$$\mathrm{NFT}(\boldsymbol{\Phi q}(t)) \rightarrow \begin{cases} \hat{a}(\lambda) = a(\lambda) \\ \hat{b_m}(\lambda) = b_m(\lambda)\mathrm{e}^{\mathrm{j}\phi_m} \\ \hat{\boldsymbol{q}}_{\boldsymbol{c}}(\lambda) = \boldsymbol{\Phi}^{-1}\boldsymbol{q}_{\mathbf{c}}(\lambda) \\ \hat{\boldsymbol{q}}_{\boldsymbol{d}}(\lambda) = \boldsymbol{\Phi}^{-1}\boldsymbol{q}_{\mathbf{d}}(\lambda_k) \end{cases}, \tag{5.20}$$

where $\boldsymbol{\Phi} = diag(\mathrm{e}^{\mathrm{j}\phi_1}, \mathrm{e}^{\mathrm{j}\phi_2}, \cdots, \mathrm{e}^{\mathrm{j}\phi_M})$ is the diagonal matrix containing the constant phase terms, $\boldsymbol{\Phi}^{-1}$ is its inverse, $\hat{a}(\lambda)$, $\hat{b_m}(\lambda)$, $\hat{\boldsymbol{q}}_{\boldsymbol{c}}(\lambda)$, $\hat{\boldsymbol{q}}_{\boldsymbol{d}}(\lambda)$ are NFT parameters related to $\hat{\boldsymbol{q}}(t) = \boldsymbol{\Phi q}(t)$ and $a(\lambda)$, $b_m(\lambda)$, $\boldsymbol{q}_{\boldsymbol{c}}(\lambda)$, $\boldsymbol{q}_{\boldsymbol{d}}(\lambda)$ are NFT parameters related to $\boldsymbol{q}(t)$. Vectors $\hat{\boldsymbol{q}}_{\boldsymbol{c}}(\lambda)$, $\hat{\boldsymbol{q}}_{\boldsymbol{d}}(\lambda)$, $\boldsymbol{q}_{\mathbf{c}}(\lambda)$ and $\boldsymbol{q}_{\mathbf{d}}(\lambda_k)$ have entries $\hat{q}_{\mathrm{c}}^{(m)}(\lambda)$, $\hat{q}_{\mathrm{d}}^{(m)}(\lambda)$, $q_{\mathrm{c}}^{(m)}(\lambda)$ and $q_{\mathrm{d}}^{(m)}(\lambda_k)$ respectively.

5.2.2. Time Shift

From the LP property, or more specifically the propagator $\boldsymbol{K}_i$ in section 5.2.4, it can be seen, that the nonlinear spectral components $a(\lambda)$, $b_m(\lambda)$, $q_c^{(m)}(\lambda)$ of a pulse-set $\boldsymbol{q}(t)$ are modified according to

$$\hat{a}(\lambda) = a(\lambda)\mathrm{e}^{\mathrm{j}\lambda t_0} \tag{5.21}$$

$$\hat{b}_m(\lambda) = b_m(\lambda)\mathrm{e}^{-\mathrm{j}\lambda t_0} \tag{5.22}$$

$$\hat{q}_{\mathrm{c}}^{(m)}(\lambda) = q_{\mathrm{c}}^{(m)}(\lambda)\mathrm{e}^{-2\mathrm{j}\lambda t_0}, \tag{5.23}$$

for a time shifted set $\hat{\boldsymbol{q}}(t) = \boldsymbol{q}(t - t_0)$ and its NFT related parameters $\hat{a}(\lambda)$, $\hat{b}_m(\lambda)$, $\hat{q}_{\mathrm{c}}^{(m)}(\lambda)$.

5.2.3. Trace Formula

The trace formula known for the scalar case can be extended as well. The unimodularity condition extends straight-forward from the scalar case as $|a(\lambda)|^2 + \sum_{m=1}^{M} |b_m(\lambda)|^2 = 1$ and will be used to extend the trace formula from the scalar case. Using the derivation from the book of Ablowitz and Segur [90] the expressions for the energy confined in

the nonlinear continuous and discrete spectrum are given by

$$E_\mathrm{c} = E(\boldsymbol{q}_\mathbf{c}(\lambda)) = \frac{1}{\pi}\int_{-\infty}^{\infty} \log\left(1 + \sum_{m=1}^{M} |q_\mathrm{c}^{(m)}(\lambda)|^2\right) \mathrm{d}\lambda \tag{5.24}$$

$$E_\mathrm{d} = 4\sum_{k=1}^{K} \Im\{\lambda_k\}. \tag{5.25}$$

5.2.4. Layer-Peeling Property

A formulation for the LP property still exists for the multi-mode case. In the following, it is assumed that two non-overlapping signals $\boldsymbol{q}_1(t)$, $\boldsymbol{q}_2(t)$, which are zero outside of intervals $[t_1, t_2]$, $[t_2, t_3]$ respectively, have nonlinear Fourier transforms $\mathrm{NFT}(\boldsymbol{q_1}(t)) \rightarrow \{a^{(1)}(\lambda), b_m^{(1)}(\lambda)\}$ and $\mathrm{NFT}(\boldsymbol{q_2}(t)) \rightarrow \{a^{(2)}(\lambda), b_m^{(2)}(\lambda)\}$. If the joint signal $\boldsymbol{q}(t) = \boldsymbol{q_1}(t) + \boldsymbol{q_2}(t)$ has nonlinear Fourier transform $\mathrm{NFT}(\boldsymbol{q}(t)) \rightarrow \{a(\lambda), b_m(\lambda)\}$, the relation between the nonlinear Fourier coefficients is given by

$$a(\lambda) = a^{(1)}(\lambda)a^{(2)}(\lambda) - \sum_{m=1}^{M} b_m^{(1)}(\lambda) b_m^{(2),*}(\lambda) \tag{5.26}$$

$$b_m(\lambda) = b_m^{(2)}(\lambda)a^{(1)}(\lambda) + b_m^{(1)}(\lambda)a^{(2),*}(\lambda). \tag{5.27}$$

The above relations can be obtained by defining the propagation operator

$$\boldsymbol{K_i} = \begin{bmatrix} a^{(i)}(\lambda) & -b_1^{(i),*}(\lambda) & -b_2^{(i),*}(\lambda) & \cdots & -b_M^{(i),*}(\lambda) \\ b_1^{(i)}(\lambda) & a^{(i),*}(\lambda) & 0 & \cdots & 0 \\ b_2^{(i)}(\lambda) & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ b_M^{(i)}(\lambda) & 0 & \cdots & 0 & a^{(i),*}(\lambda) \end{bmatrix} \tag{5.28}$$

and using

$$[a(\lambda)^{(1)}, b_1(\lambda)^{(1)}, \cdots, b_M(\lambda)^{(1)}]^\mathrm{T} = \boldsymbol{K_1}[1, 0, \cdots, 0]^\mathrm{T} \tag{5.29}$$

$$[a(\lambda)^{(2)}, b_1(\lambda)^{(2)}, \cdots, b_M(\lambda)^{(2)}]^\mathrm{T} = \boldsymbol{K_2}[1, 0, \cdots, 0]^\mathrm{T} \tag{5.30}$$

and the linearity of the propagation of eigenvectors in t. Computing $\boldsymbol{K} = \boldsymbol{K_2}\boldsymbol{K_1}[1, 0, \cdots, 0]^\mathrm{T}$ then yields the expressions for the nonlinear Fourier coefficients of the joint signal $\boldsymbol{q}(t)$ The LP property can be used to e.g. implement the LP-NFT, as presented in section

5.4.1.

5.2.5. Propagation of Nonlinear Quantities

The limit $\lim_{t\to\pm\infty} \mathbf{M}(z)$ for the operator defined in Eqs. (5.3) to (5.7) is reduced to a diagonal matrix with entries

$$M_{kk} = \begin{cases} \mathrm{j}2\lambda^2 & \text{for } k = 1 \\ -\mathrm{j}2\lambda^2 & \text{else} \end{cases} . \tag{5.31}$$

This is due to $\boldsymbol{q}(t) \to \mathbf{0}$ for $t \to \pm\infty$. Using this limit and the boundary conditions in Eq. (5.8), the derivation for the transfer functions for the nonlinear spectra is done analogous to the scalar case [43] and results in

$$\boldsymbol{q}_\mathrm{c}(\lambda, z) = \boldsymbol{q}_\mathrm{c}(\lambda, 0)\mathrm{e}^{-4\mathrm{j}\lambda^2 z} \tag{5.32}$$

$$\boldsymbol{q}_\mathrm{d}(\lambda_k, z) = \boldsymbol{q}_\mathrm{d}(\lambda_k, 0)\mathrm{e}^{-4\mathrm{j}\lambda_k^2 z}. \tag{5.33}$$

The discrete eigenvalues remain unchanged during propagation in z.

5.3. Special Solutions

In general, the NFT of most pulseshapes has to be computed numerically. However, for the NLSE-case, some well known pulseshapes have some closed form results for their nonlinear Fourier spectra (see section 4.3). For the presented multi-mode NFT, two sets of pulses are studied analytically and some closed form results on their nonlinear Fourier spectra and other properties are given. Those two pulse-sets are, the rectangular pulse-set with identical support, and the vectorial fundamental soliton (VFS).

5.3.1. The Rectangular Pulse-Set

For the rectangular pulse-set several NFT related quantities can be obtained analytically. It is assumed that, while the amplitudes A_m of the pulses in the set can be different for each mode, all pulses in the set have the same temporal support $[t_1, t_2]$.

This set of rectangular pulses is given by

$$q_m(t) = \begin{cases} A_m & -\frac{T}{2} < t < \frac{T}{2} \\ 0 & \text{else} \end{cases} . \tag{5.34}$$

Here it is assumed, that the pulse-set is zero-centered with $t_2 = -t_1 = \frac{T}{2}$. The results obtained in the following can be generalized using the relations in section 5.2.2.

Obtaining the nonlinear Fourier coefficients in closed form is done similarly to the NLSE-case in section 4.3.1. Nonetheless, a brief overview of the proofs for the following results is given in appendix A.6. The nonlinear Fourier coefficients for the continuous spectrum are obtained by

$$a(\lambda) = \left(\cos(\Delta T) - \mathrm{j}\frac{\lambda}{\Delta}\sin(\Delta T)\right) \mathrm{e}^{\mathrm{j}\lambda(t_2-t_1)} \tag{5.35}$$

$$b_m(\lambda) = -\frac{A_m^*}{\Delta}\sin(\Delta T)\mathrm{e}^{-\mathrm{j}\lambda(t_1+t_2)} \tag{5.36}$$

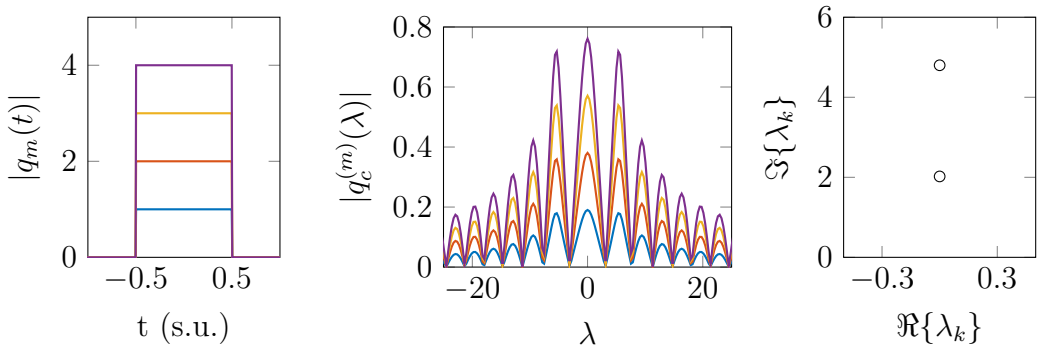
where $\Delta = \sqrt{\lambda^2 + \sum_{m=1}^{M} |A_m|^2}$. Further, the discrete eigenvalues fulfill

$$\mathrm{j}\tan\left(T\sqrt{\sum_{m=1}^{M}|A_m|^2 + \lambda^2}\right) = \sqrt{1 + \frac{\sum_{m=1}^{M}|A_m|^2}{\lambda^2}}. \tag{5.37}$$

The discrete nonlinear spectral amplitudes can be found by using the given closed form expression for $a(\lambda)$ and deriving $a'(\lambda)$ from it. An example of the NFT of such a set of rectangular pulses is given in Fig. 5.1. Note, that continuous spectral quantity $q_{\mathrm{c}}^{(m)}(\lambda)$ does not only correspond to the rectangular pulse with amplitude A_m. Due to nonlinear coupling in the fiber, every pulse in the pulse-set influences the shape of every entry in $\boldsymbol{q}_{\mathbf{c}}(\lambda)$.

5.3.2. The Fundamental Soliton

Another pulse-set for which several quantities can be obtained analytically, which also is considered a special solution of Eq. (5.1), is the VFS. It is the multi-mode analog to the fundamental soliton, described in section 4.3.2 for the NLSE case. It can be obtained, using the first iteration of the multi-mode DT, which will be presented in



(a) Time Domain (b) Continuous Spectrum (c) Discrete Eigenvalues

Figure 5.1.: NFT of a set of rectangular pulses

section 5.4.3. The VFS pulse-set is given by

$$q_m(t,z) = -\mathrm{j}4\eta \cdot \mathrm{e}^{-\mathrm{j}\phi_m} \cdot \mathrm{e}^{-\mathrm{j}4(\xi^2-\eta^2)z} \cdot \mathrm{e}^{-\mathrm{j}2\xi t} \cdot \frac{\mathrm{e}^{-8\xi\eta z}\mathrm{e}^{-2\eta t}\frac{|q_\mathrm{d}^{(m)}(\lambda_k)|}{2\eta}}{1+\mathrm{e}^{-16\xi\eta z}\mathrm{e}^{-4\eta t}\sum\limits_{s=1}^{M}\left(\frac{|q_\mathrm{d}^{(s)}(\lambda_k)|}{2\eta}\right)^2}, \tag{5.38}$$

where the discrete spectral amplitudes are $q_\mathrm{d}^{(m)}(\lambda_k) = |q_\mathrm{d}^{(m)}(\lambda_k)|\mathrm{e}^{\mathrm{j}\phi_m}$ and $\lambda_k = \xi + \mathrm{j}\eta$. The continuous spectrum of this pulse is $\boldsymbol{q_\mathrm{c}}(\lambda) = \mathbf{0}$ for all $\lambda \in \mathbb{R}$. It can be seen, that each single pulse $q_m(t,z)$ now no longer has the ideal sech$(\cdot)$ shape that the NLSE solitons have. For the VFS pulse-set, there are further quantities that can be computed analytically. In the following, closed-form expressions for temporal width, linear bandwidth and time-bandwidth product are presented for the general M-mode VFS pulse-set given by Eq. (5.38).

Temporal Width

For the VFS, the temporal width can be computed analytically. Similar to the treatment in section 4.3.2, the energy of the pulse-set can be related to its width by

$$E_\mathrm{vfs} \cdot p = \sum_{m=1}^{M} \int_{-T/2+t_0}^{T/2+t_0} |q_m(t,z)|^2 \mathrm{d}t, \tag{5.39}$$

where t_0 is the pulse-sets center and E_vfs is the total energy of the VFS. There is a certain intricacy linked to setting a meaningful value for t_0, since pulse-position as well

as the energy of the components $q_m(t,z)$ should be considered. In the following, it is assumed that all pulses are centered at the same position t_0 and thus the exact position of the pulse-set can be neglected for the computation of the temporal width. This can be done by choosing the right constellation for the modulated $b_m(\lambda_k)$ values. Note, that the results using this simplification are less useful when the TBR of a system is considered instead of the TBP of a single pulse-set. By inserting Eq. (5.38) into (5.39), it can be seen that the temporal width is depending on the in-window energy percentage p and is given by

$$T = \frac{\operatorname{arctanh}(p)}{\eta}. \tag{5.40}$$

This result is identical to the result for the NLSE case given in section 4.3.2. A short explanation of the steps taken to obtain Eq. (5.40) from Eq. (5.39) is given in appendix A.7.1.

Bandwidth

Similarly the linear bandwith can be obtained by evaluating

$$E_{\text{vfs}} \cdot p = \sum_{m=1}^{M} \int_{-B/2+f_0}^{B/2+f_0} |\mathcal{F}\{q_m(t,z)\}|^2 \mathrm{d}f, \tag{5.41}$$

using Eq. (5.38). The parameter f_0 is once again chosen such, that the integrated expression is centered on the frequency support during evaluation. Again, the intricacies of choosing a suitable value f_0 are not treated in detail, assuming that pulses are generated such that each pulse in the set has the identical value for f_0. The linear bandwidth for some given in-window energy percentage p is then given by

$$B = \frac{4\eta}{\pi^2}\operatorname{arctanh}(p). \tag{5.42}$$

This expression is once again identical to the NLSE case. The steps taken from Eq. (5.41) to (5.42) are detailed in appendix A.7.2.

Time-Bandwidth Product

The resulting TBP for the VFS

$$\mathrm{TBP}(p) = \frac{4}{\pi^2}\mathrm{arctanh}^2(p) \tag{5.43}$$

is identical to the NLSE case from section 4.3.2 and thus also is independent of the specific value of η. Note, that while the TBP as stated in Eq. (5.43) is not dependent on parameters ξ and z, the pulse-centers in time and linear frequency domain in general will shift due to changes in these two parameters. Also note, that according to Eqs. (5.43) and (4.67) there is no additional time-bandwidth penalty for VFSs. They thus, under the assumption that both NLSE and multi-mode channels are ideal, permit higher SEs by a factor M. This however comes at the price of operating a multi-mode fiber system for transmission and regarding ASE noise, also comes with the need for higher total signal-power to maintain the same SNR.

5.4. Algorithms

From the mathematical foundations given in section 5.1, algorithms for the forward and inverse NFTs can be obtained. Note, that the algorithms given here are mostly extensions of algorithms from [23], [27]. Some of the extensions exhibit potential for optimization in terms of (computational) complexity. This potential was not further explored and, for now, their performance is only considered in regards to accuracy of the performed transformation. This evaluation was done for specific signals with nonlinear spectral components that are known in closed-form (see section 5.3).

Most of the following algorithms are implemented starting from the relation $\mathbf{L}(z)\boldsymbol{v}(t,\lambda) = \lambda\boldsymbol{v}(t,\lambda)$ rewritten as $\boldsymbol{v_t}(t,\lambda) = \mathbf{P}(z)\boldsymbol{v}(t,\lambda)$ with

$$\mathbf{P}(z) = \begin{bmatrix} -\mathrm{j}\lambda & q_1 & q_2 & \cdots & q_M \\ -q_1^* & \mathrm{j}\lambda & 0 & \cdots & 0 \\ -q_2^* & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -q_M^* & 0 & \cdots & 0 & \mathrm{j}\lambda \end{bmatrix}, \tag{5.44}$$

which can be obtained by algebraic manipulation of the eigenvalue equation. Note that

$\boldsymbol{v_t}(t,\lambda) = \frac{\partial \boldsymbol{v}(t,\lambda)}{\partial t}$ still denotes the derivative in t and distance parameter z assumed to be fixed and thus omitted in the following.

For the forward transformation Eq. (5.44) is discretized and then either a variety of numerical integration schemes or a matrix eigenvalue calculation can be performed. Similar to the NLSE case, the forward NFT is computing nonlinear Fourier coefficients, based on the pulse-set in time-domain. The transformation now incorporates the signals in all stimulated modes of the multi-mode channel. The resulting nonlinear Fourier coefficients are subsequently used to compute the nonlinear Fourier transform according to Eq. (5.11).

5.4.1. Forward Transformations for the Continuous Spectrum

In a first step, algorithms to obtain the continuous nonlinear Fourier spectrum will be treated. While there are many algorithms that can be modified to work in the multi-mode case, the following section is only presenting two methods, either because they are deemed especially instructive or because they yield high accuracy results. The presented methods are, the FD scheme which is used as an instructive example and the LP method, which is extended to the multi-mode case since it showed good accuracy for moderate complexity in the NLSE case.

Forward Discretization Scheme

This scheme is known to have low accuracy and rather high complexity, but is simple in structure and will thus serve as an introductory example. The signal will be discretized on a temporal grid identical to the scalar case (see section 4.4). The matrix $\mathbf{P}$ in $\boldsymbol{v_t}(t,\lambda) = \mathbf{P}\boldsymbol{v}(t,\lambda)$ is assumed to be constant for each step h and thus $\boldsymbol{v}[n+1] = (\boldsymbol{I}_{M+1} + h\boldsymbol{P}[n])\boldsymbol{v}[n]$ with (compare forward Euler-scheme used in Eq. (4.71))

$$\boldsymbol{P}[n] = \begin{bmatrix} -\mathrm{j}\lambda & q_1[n] & q_2[n] & \cdots & q_M[n] \\ -q_1^*[n] & \mathrm{j}\lambda & 0 & \cdots & 0 \\ -q_2^*[n] & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -q_M^*[n] & 0 & \cdots & 0 & \mathrm{j}\lambda \end{bmatrix}. \tag{5.45}$$

The initial condition for $\boldsymbol{v}[0]$ as

$$\boldsymbol{v}[0] = [1, 0, \cdots, 0]^{\mathrm{T}} \mathrm{e}^{-\mathrm{j}\lambda t_1}. \tag{5.46}$$

Here $t_1 = t[0]$ is the temporal position of samples $q_m[0]$, $\boldsymbol{I}_{M+1}$ is the $(M+1)\mathrm{x}(M+1)$ identity matrix and the initial condition vector $\boldsymbol{v}[0]$ has $M+1$ entries. With this, one can iterate $\boldsymbol{v}[0]$ towards $\boldsymbol{v}[N_\mathrm{s}-1]$ and then obtain the nonlinear Fourier coefficients from

$$\boldsymbol{v}[N_\mathrm{s}-1] \approx \begin{bmatrix} a(\lambda)\mathrm{e}^{-\mathrm{j}\lambda t_2} \\ b_1(\lambda)\mathrm{e}^{\mathrm{j}\lambda t_2} \\ \vdots \\ b_M(\lambda)\mathrm{e}^{\mathrm{j}\lambda t_2} \end{bmatrix} \tag{5.47}$$

where $t_2 = t[N_\mathrm{s}-1]$ is the last value of the discrete temporal support.

Layer Peeling Scheme

The LP method for the NLSE case did show good results and was thus conjectured to also prove more stable for the multi-mode case than e.g. the extension of the rather naive FD scheme. The complexity of the FD and LP method are comparable for the multi-mode case as well.

For this method, the used parameters are initialized as $a[0] = 1$ and $b_m[0] = 0 \,\forall\, m = 1, \cdots, M$. The update step can be written as

$$a[n+1] = a[n]x[n] - \sum_{m=1}^{M} b_m[n] y_m^*[n] \tag{5.48}$$

$$b_m[n+1] = a[n]y_m[n] + b_m[n]x^*[n] \tag{5.49}$$

with the update parameters $x[n]$ and $y_m[n]$ given by the closed form expression of the nonlinear Fourier parameters of a rectangular pulse of width $h_n = h[n] = t[n] - t[n-1]$

and amplitudes $q_m[n]$. They are given by

$$\Delta = \sqrt{\lambda^2 + \sum_{m=1}^{M} |q_m[n]|^2} \tag{5.50}$$

$$x[n] = (\cos(\Delta h_n) - \mathrm{j}\lambda \mathrm{si}(\Delta h_n) h_n)\, \mathrm{e}^{\mathrm{j}\lambda(t[n]-t[n-1])} \tag{5.51}$$

$$y_m[n] = -\, q_m[n]^* \mathrm{si}(\Delta h_n) h_n \mathrm{e}^{-\mathrm{j}\lambda(t[n]+t[n+1])}. \tag{5.52}$$

The nonlinear Fourier coefficients of the transformed pulse-set are then obtained as $a(\lambda) \approx a[N_\mathrm{s} - 1]$ and $b_m(\lambda) \approx b_m[N_\mathrm{s} - 1]$. A sketch of the derivation, which is similar to the derivation for the NLSE case, is given in appendix A.8.

5.4.2. Forward Transformations for the Discrete Spectrum

Some methods to compute parameters of the discrete nonlinear Fourier spectrum, presented in section 4.4.2, can be extended to the multi-mode case as well. In the following, some algorithms will be presented, that either are instructive or exhibit beneficial properties in regards to complexity and accuracy.

Forward Discretization Scheme

Analog to the NLSE case, many of the numerical integration based schemes can be extended to deal with the discrete spectrum via e.g. the Newton-Raphson search algorithm, described in section 4.4.2. The extension for the FD scheme given section 5.4.1 is obtained using the same steps as in the NLSE case.

The derivative of $\boldsymbol{v}[n+1] = (\boldsymbol{I}_{M+1} + h\boldsymbol{P}[n])\boldsymbol{v}[n]$ in λ which is

$$\boldsymbol{v}'[n+1] = (h\boldsymbol{P}'[n])\boldsymbol{v}[n] + (h\boldsymbol{P}[n])\boldsymbol{v}'[n], \tag{5.53}$$

with

$$\boldsymbol{P}'[n] = \begin{bmatrix} -\mathrm{j} & 0 & 0 & 0 \\ 0 & \mathrm{j} & 0 & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \mathrm{j} \end{bmatrix}, \tag{5.54}$$

can be used to update the derivatives of the eigenvectors. The initial values are given by

$$\boldsymbol{v}'[0] = [-\mathrm{j}\lambda, 0, \cdots, 0]^\mathrm{T} \mathrm{e}^{-\mathrm{j}\lambda t_1} \tag{5.55}$$

and the derivative $a'(\lambda)$ can be obtained using

$$a'(\lambda) \approx v_1'[N_\mathrm{s}-1]\mathrm{e}^{\mathrm{j}\lambda t_2} + v_1[N_\mathrm{s}-1]\mathrm{j}\lambda\mathrm{e}^{\mathrm{j}\lambda t_2}, \tag{5.56}$$

where v_1 is the first entry of $\boldsymbol{v}$ and v_1' is its derivative in λ. With these equations, the search algorithm from [27] can be used to obtain the discrete eigenvalues and the corresponding spectral amplitudes $\boldsymbol{q}_\mathbf{d}(\lambda_k)$ or $b_m(\lambda_k)$-values.

Layer-Peeling Scheme

The LP method is another integration based scheme, which already was extended for the continuous spectrum in section 5.4.1. It can be extended for the detection of discrete spectral components as well. To compute the additional parameter $a'(\lambda_k)$ (and $b_m'(\lambda_k)$), the set of equations

$$\begin{aligned} a'[n+1] =& a'[n]x[n] + a[n]x'[n] \\ &- \left(\sum_{m=1}^{M} b_m'[n]y_m^*[n] + \sum_{m=1}^{M} b_m[n]y_m^{*\prime}[n]\right) \end{aligned} \tag{5.57}$$

$$b_m'[n+1] = a'[n]y_m[n] + a[n]y_m'[n] + b_m'[n]x^*[n] + b_m[n]x^{*\prime}[n] \tag{5.58}$$

can be used. These equations are the λ derivatives of Eqs. (5.48) and(5.49). The update values $x'[n]$, $y_m'[n]$ are

$$x'[n] = \mathrm{j}h_n(1-\frac{\lambda^2}{\Delta^2})(\cos(\Delta h_n) - \frac{\sin(\Delta h_n)}{\Delta h_n})\mathrm{e}^{\mathrm{j}\lambda h_n} \tag{5.59}$$

$$\begin{aligned} y_m'[n] =& -q_m^*[n]\left[\frac{\lambda h_n}{\Delta^2}\cos(\Delta h_n)\right. \\ &\left. - \left(\frac{\lambda}{\Delta^3} + \mathrm{j}\frac{t_n+t_{n-1}}{\Delta}\right)\sin(\Delta h_n)\right]\mathrm{e}^{-\mathrm{j}\lambda(t_n+t_{n-1})} \end{aligned} \tag{5.60}$$

Note that Eqs. (5.59) and (5.60) are identical to the NLSE case equations in [27], except for the modified equation for Δ, which is given in Eq. (5.50). Some short notes on the derivation of the relations above are given in appendix A.8. In Fig. 5.2 the modulated $b_m(\lambda_k)$-values of a vectorial multi-soliton pulse ($M = 4$) are shown for B2B transmission and transmission over $\mathcal{L} = 100$km of the ideal multi-mode channel for generation and detection by FD and LP respectively. Compensation of the channel perturbations was done using Eq. (5.33). The accuracy of both methods differ by

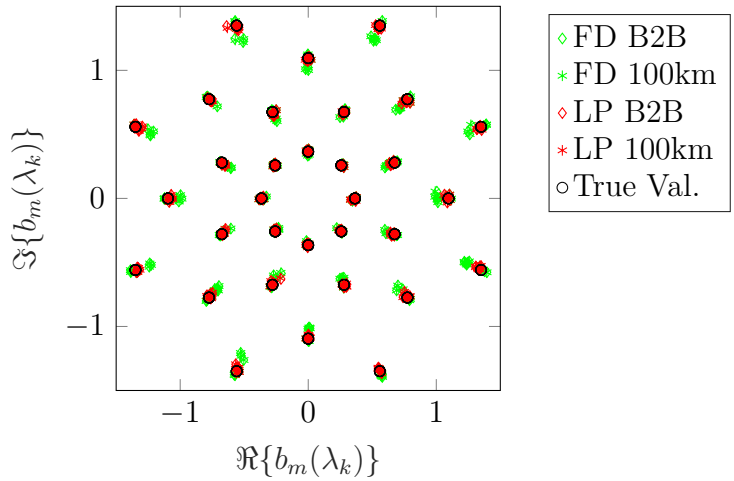


Figure 5.2.: $b_m(\lambda_k)$-values of a multi-soliton with $K = 3$ discrete eigenvalues

approximately one order of magnitude. While, depending on the constellation used, FD might infer a non-negligible reduction in detection accuracy, for LP at least for constellations with moderately high cardinality the error can be assumed to be negligible compared to impairments by either noise or nonlinear distortions in systems with nonzero fiber loss. Note that this is only meant to be a small scale test of both methods as $\mathcal{L} = 100$km channel length is not in the region studied primarily in this thesis.

Matrix Eigenvalue Methods

Finding discrete eigenvalues for the multi-mode case can also be facilitated by solving a matrix-eigenvalue problem. There is a variety of algorithms, e.g. in [27], that can be extended. For purposes of demonstrating the approach, the CDEP from section 4.4.2 is used. We can state the extended discretized eigenvalue problem as

$$\hat{\boldsymbol{L}}\hat{\boldsymbol{v}} = \lambda\hat{\boldsymbol{v}} \tag{5.61}$$

where $\hat{\boldsymbol{v}}$ is the concatenation of all vectors $\boldsymbol{v_m} = [v_m[1], v_m[2], \cdots, v_m[N_\mathrm{s}]]^\mathrm{T}$, such that $\hat{\boldsymbol{v}}$ has length $N_\mathrm{s} \cdot (M+1)$. The matrix $\hat{\boldsymbol{L}}$ then has the form

$$\hat{\boldsymbol{L}} = \begin{bmatrix} \boldsymbol{D} & -\boldsymbol{Q_1} & -\boldsymbol{Q_2} & \cdots & -\boldsymbol{Q}_M \\ -\boldsymbol{Q}_1^* & \text{-}\boldsymbol{D} & \boldsymbol{0}_{N_\mathrm{s}} & \cdots & \boldsymbol{0}_{N_\mathrm{s}} \\ -\boldsymbol{Q}_2^* & \boldsymbol{0}_{N_\mathrm{s}} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \boldsymbol{0}_{N_\mathrm{s}} \\ -\boldsymbol{Q}_M^* & \boldsymbol{0}_{N_\mathrm{s}} & \cdots & \boldsymbol{0}_{N_\mathrm{s}} & \text{-}\boldsymbol{D} \end{bmatrix} \tag{5.62}$$

where $\boldsymbol{0}_{N_\mathrm{s}}$ is the all-zero matrix of size N_sxN_s, the matrix $\boldsymbol{Q_m} = \mathrm{diag}([q_m[1], q_m[2]], \cdots, q_m[M])$ and $\boldsymbol{D}$ is the central finite difference matrix of size N_sxN_s

$$\boldsymbol{D} = \frac{1}{2h} \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & -1 \\ -1 & 0 & 1 & \ddots & & 0 \\ 0 & -1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 & 0 \\ 0 & & \ddots & -1 & 0 & 1 \\ 1 & 0 & \cdots & 0 & -1 & 0 \end{bmatrix} \tag{5.63}$$

The discrete eigenvalues are then obtained by finding the eigenvalues of $\hat{\boldsymbol{L}}$. Note that, if a general purpose eigenvalue solver is used the complexity of this method is up to $\mathcal{O}((N_\mathrm{s} \cdot (M+1))^3)$, depending on the solvers exact implementation. This method also (at least for the central difference version) is not very accurate. Additionally, this method only gives the discrete eigenvalues. Thus, another method has to be used additionally to obtain the spectral coefficients. It can be seen, that in most cases this matrix-based approach is not favorable for setups in which guesses for positions of discrete eigenvalues exist. There are other types of matrix-methods described in [27], that can be extended similarly, that might give a higher accuracy, however the complexity can only be reduced by finding a more effective way of solving the matrix eigenvalue problem itself.

5.4.3. Inverse NFT for the Discrete Spectrum

For purely discrete pulses the DT can be extended for the multi-mode case. In fact, the method was already described in a suitable manner in [122] and extended to the

two-polarization case in [143]. Nonetheless, the method is shortly reviewed in the following, using the notation chosen for this thesis. In a first step, vectors

$$\boldsymbol{v}(\lambda_k, t) = [C_1(\lambda_k)\mathrm{e}^{-\mathrm{j}\lambda_k t}, C_2(\lambda_k)\mathrm{e}^{\mathrm{j}\lambda_k t}, C_3(\lambda_k)\mathrm{e}^{\mathrm{j}\lambda_k t}, \cdots, C_{M+1}(\lambda_k)\mathrm{e}^{\mathrm{j}\lambda_k t}], \tag{5.64}$$

are defined. They are solutions of $\boldsymbol{v_t}(\lambda_k, t) = \boldsymbol{P}\boldsymbol{v}(\lambda_k, t)$ for the zero-pulse-set $\boldsymbol{q}(t) = \boldsymbol{0}$, with some parameters $C_m(\lambda_k)$ for $k = 1, \cdots, K$ and $m = 1, \cdots, M+1$. These vectors can be used to update the trivial time domain solution $\boldsymbol{q}(t) = \boldsymbol{0}$, by adding some discrete eigenvalue λ_0 to its discrete nonlinear Fourier spectrum. This is done by

$$\hat{q}_m^{(i)}(t) = q_m^{(i)}(t) + 2\mathrm{j}(\lambda_0^* - \lambda_0)\frac{v_1(\lambda_0, t)v_{m+1}(\lambda_0, t)^*}{|v_1(\lambda_0, t)|^2 + \sum\limits_{s=1}^{M}|v_{m+1}(\lambda_0, t)|}. \tag{5.65}$$

This relation can be used in an iterative manner, to add multiple discrete eigenvalues to the discrete spectrum of a pulse-set, but the eigenvectors have to be updated after each addition of an discrete eigenvalue to the nonlinear spectrum. To facilitate this, [122] provides an update method. The new eigenvector $\hat{\boldsymbol{v}}_{\boldsymbol{t}}(\lambda_k, t) = \hat{\boldsymbol{P}}\hat{\boldsymbol{v}}(\lambda_k, t)$ for $\hat{\boldsymbol{P}}$, now parametrized by $\hat{\boldsymbol{q}}(t)$. The pulse-set $\hat{\boldsymbol{q}}(t)$ has a discrete nonlinear Fourier spectrum, which contains the discrete eigenvalue λ_0. The updated vectors are obtained by

$$\hat{\boldsymbol{v}}(\lambda_k, t) = (\lambda_k \boldsymbol{I} - \boldsymbol{G_0})\boldsymbol{v}(\lambda_k, t) \tag{5.66}$$

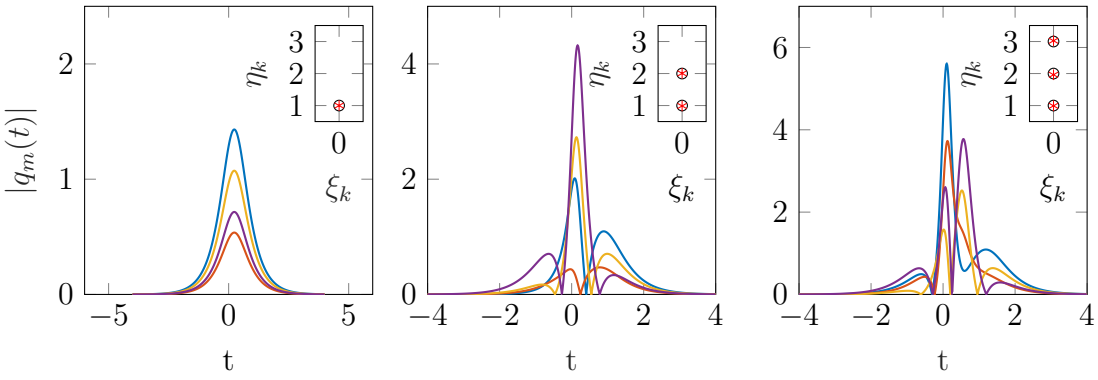
with

$$\boldsymbol{G_0} = \boldsymbol{\Psi}\boldsymbol{M_0}\boldsymbol{\Psi^{-1}} \text{ where } \boldsymbol{M_0} = \mathrm{diag}(\lambda_0, \lambda_0^*, \lambda_0^*, \cdots, \lambda_0^*) \tag{5.67}$$

and

$$\boldsymbol{\Psi} = \begin{bmatrix} v_1(\lambda_0, t) & v_2^*(\lambda_0, t) & \cdots & v_{M+1}^*(\lambda_0, t) \\ v_2(\lambda_0, t) & -v_1^*(\lambda_0, t) & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ v_{M+1}(\lambda_0, t) & 0 & \cdots & -v_1^*(\lambda_0, t) \end{bmatrix}. \tag{5.68}$$

It can be seen, that this enables the generation of multi-soliton pulse-sets for M modes by iterating over the steps outlined above. In Fig. 5.3, the stepwise generation of such a vector-multisoliton is shown, together with its expected (black rings), as well as its numerically recovered (red asterisks) discrete spectrum in a B2B setup. For estimation of the discrete eigenvalues in Fig. 5.3, the FD algorithm, described in section 5.4.1 was used. While the single per-mode signals do no longer have the shape of soliton



(a) $\{\lambda_k\} = \{\mathrm{j}\}$ (b) $\{\lambda_k\} = \{\mathrm{j}, 2\mathrm{j}\}$ (c) $\{\lambda_k\} = \{\mathrm{j}, 2\mathrm{j}, 3\mathrm{j}\}$

Figure 5.3.: Multi-soliton generation by the DT ($\lambda_k = \xi + \mathrm{j}\eta$)

pulses and thus, if propagated in isolation, do not exhibit the behavior associated with solitons, the jointly propagating pulse-set retains its solitonic behavior, due to nonlinear coupling between all co-propagating stimulated modes.

6
Conclusions

In this thesis, several aspects and use-cases of the NFT for fiber optic communication systems have been studied. In doing so, several techniques for single polarization Schrödinger equation were tested, ranging from advanced transformation algorithms and alternative receiver structures to measurement methods for channel parameters utilizing the nonlinear Fourier spectrum. Building on the work on dual-polarization NFTs published in recent years, many of the relations and algorithms obtained for the single polarization channel, were extended to a channel model for the strong coupling multi mode case. This chapter aims to give a compressed overview of the most important treated topics and results, as well as some open questions that could be starting points for future work in the field. In concluding this chapter, a short section with acknowledgements is included.

6.1. Overview

In a first step, the theoretical fundamentals for the NFT and its inverse were studied in great detail. This includes the study of the special properties of the quantities in the nonlinear Fourier domain, e.g. in regards to temporal shifts or spatial propagation along the channel. One of the main contributions of the thesis in this field is a theoretical treatment of the time-bandwidth product and its discussion in light of transmission system design for data transmission via purely solitonic pulses.

After presenting the basic underlying concepts for the NFT, a wide range of algorithms for forward and backward transformations was presented alongside instructions on how to implement them in numerical simulations. They were tested in terms of their accuracy in detecting the different components of the nonlinear Fourier spectrum. In some cases, studies regarding the computational complexity were done as well. They were carried out either analytically or by using measurements of execution time. One of the main contributions in this field, was the extension of the eigenvalue removal method for full spectrum pulses and its subsequent evaluation.

Since there is only scarce theoretical analysis of the influence of perturbations to the channel model used for the derivation of the NFT, the effect of these perturbations, in particular the influence of fiber-loss on nonlinear Fourier domain quantities, was studied. The path-loss average technique to stabilize solitonic pulses during transmission in a lossy medium was studied and evaluated numerically for several transmission scenarios. Also, the trajectories of discrete eigenvalues during propagation in a lossy medium and the influence of modulation on these trajectories were studied numerically. Similarly, phenomena like "absorption" of a discrete eigenvalue into the real axis and "collision" of two discrete eigenvalues in the complex plane were studied.

Another field in which several specific topics were studied, is the improvement of NFT-aided transmission by additional (digital) signal processing. This includes the developed and tested labeling schemes for on-off keying soliton transmission, which lead to strongly decreased bit-error per symbol-error ratios for moderate transmission distances. An optimum constellation for transmission under timing-jitter as the only perturbation was derived. In another scheme, matched-filters were used for the detection of solitonic pulses. This scheme in particular improved significantly on the tested detection algorithms and gave some deeper insights into problems regarding noise in NFT-aided systems. To improve detection at the receiver in an additional digital signal processing step, some well known clustering algorithms were tested, showing some potential to improve the bit-error rate of the tested systems.

Outside of the topics directly related to transmission schemes, parameter measurement for optical fibers, using either fundamental solitons or multi-solitons was studied. Apart from showing that these types of measurement schemes in general can be used to obtain the nonlinearity and dispersion parameter of a tested fiber, the results also indicated, that rescaling of NFT generated pulses, using the path-loss average model studied earlier, was not optimal. It was shown, that by using the nonlinearity measurements from the parameter estimation for rescaling the pulses prior to transmitting

them, data transmission can generally be improved. This is due to the fact, that the measurement algorithms tested are estimating the physical nonlinearity parameter of the fiber by effectively finding the nonlinearity value for which the solitonic signals propagate most optimally. For lossy channels those two quantities are not necessarily the same.

While all of the prior studies in this thesis were conducted assuming the scalar nonlinear Schrödinger equation as the underlying channel model, the NFT for the strong coupling Manakov equation was used as an underlying channel model as well. The mathematical basics necessary for the transformations were extended, if possible, to enable the construction of a forward and backward transform for this new channel model. Additionally, analogs to several properties for the scalar channel NFT were derived and studied, as well as the derivation of several closed form solutions for special pulse-sets. Algorithms for forward and inverse transformations of continuous and discrete spectrum pulses were developed from their scalar counterparts as well.

6.2. Future Work

Even though, the interest in topics linked to NFT-aided transmission has declined in recent years, there are still many interesting open topics that have the potential to improve on the current status of NFT related theory and applications.

While there are several publications that treat the influence of white Gaussian time-domain noise on the nonlinear Fourier domain by means of perturbation theory, its influence is not fully quantified yet. Similarly, there is little knowledge on the influence of fiber-loss on the nonlinear Fourier domain up to now. Since noise and fiber-loss are perturbations that cause major problems when designing NFT-aided systems, a more complete quantification of the influence of these two quantities should be pursued.

Further, there also is no analytic relation for the width in time and linear frequency to the modulation of quantities in the nonlinear Fourier domain. This also complicates the design of systems and is, in the eyes of the author, also an important topic for future work.

As already noted in the section before, there are several avenues for improvement of NFT-aided transmissions systems that are currently not fully explored. This includes e.g. the use of the results from fiber-parameter estimation to rescale pulses for a more optimal propagation in the optical fiber channel. Similarly, as seen in the section on

the time-bandwidth requirements of multi-soliton systems, there might be a way to decrease the necessary bandwidth and time-slot in multi-soliton systems by choosing a subset of transmission pulses from a larger alphabet, which has beneficial qualities regarding its time-bandwidth requirements. The use of matched filter-banks for the detection of solitonic pulses also was briefly studied in this thesis. It can be seen, that even though complexity issues might arise quickly, the stability against noise can be increased significantly using this method. While the improvement of such a detection scheme is conjectured to be less pronounced for higher transmission powers, there still might be some room for improvement for certain types of transmission systems.

With the currently growing interest in space division multiplexing to scale the achievable throughput of currently used transmission schemes, further studies of NFTs for new channel models could also be reasonably chosen as a field for future studies. As could be seen in the section on the NFT for the strong-coupling Manakov equation, there are still several properties and algorithms known for the scalar case, that are not readily available in the mulit-mode case. In particular, there is no inverse transformation for the continuous spectrum presented in this thesis. An extension of the inverse Ablowitz-Ladik discrete layer-peeling algorithm was studied, but there seemed to be no viable possibility to extend it for the multi-mode case. There are several other methods for the inverse continuous spectrum transformation known for the scalar case that could be studied in terms of their applicability to this new channel. Similarly, there currently is no known eigenvalue-removal method, forward-backwards transformation or full spectrum inverse NFT.

Another topic for future work could be the study of the influence of perturbations like e.g. noise and fiber-loss on the nonlinear spectra in the multi-mode case. Over the course of the studies that lead to this thesis, an attempt was made to get some insight into the noise characteristics in the nonlinear Fourier domain, assuming white noise in time-domain. While a perturbation similar to the way it is done in literature for the scalar case was attempted, it seems that the left-hand eigenvector of the Lax operator, readily available in the scalar case, could not be obtained in a similar fashion. If this quantity would be obtainable, we conjecture, that the remaining part of the perturbation approach would lead to first-order noise statistics in a rather straightforward manner.

The theoretical framework of the NFT restricts the construction of transformations to channel models with certain characteristics. Still, an interesting avenue for further studies in the field could be the derivation of NFT transformations for more general

channel models linked to the field of space division multiplexing.

6.3. Acknowledgement

In concluding the main part of this thesis, the author would like to express his gratitude for the help of several people. First and foremost the author would like to thank professor Hanik for the opportunity to write this thesis as well as the helpful comments and discussions throughout the last 7 years. Similarly, the author would like to express his gratitude towards his former colleagues at the institute, including scientific and non-scientific staff. In particular the author would like to thank Javier Garcia and Paolo Carniello for the support and the friendly atmosphere in the shared office. Thank you to all my collaborators, especially Bensu Baran, Daniel Plabst and Thomas Göttsberger. Finally, the author would like to thank his partner Sabrina for the emotional support in the home stretch of finishing this doctoral thesis.

A

Additional Proofs

The following proofs were moved to the appendix from their respective sections to improve the readability and structure of the thesis. They are referenced in the main text and use equations given in the main chapters as starting points. Note that in some of the following sections the notation is shortened by e.g. omitting unnecessary indexing or function dependencies, to improve the readability. If not straightforward, omissions and simplifications will be explicitly stated in the sections below. With exception of the two fundamental proofs given in sections A.1 and A.2, the proofs in this section are not taken directly from references in literature. They thus can be considered to be results of this thesis.

A.1. The Lax Equation

Starting from Eq. (4.7) and omitting spatial and temporal dependencies in the notation, the derivative in z is given by

$$\begin{aligned} \mathrm{L}_z &= \frac{\mathrm{d}}{\mathrm{d}z}(\mathrm{G}\Lambda\mathrm{G}^{-1}) \\ &= \mathrm{G}_z\Lambda\mathrm{G}^{-1} + \mathrm{G}\Lambda(-\mathrm{G}^{-1}\mathrm{G}_z\mathrm{G}^{-1}). \end{aligned}$$

Further $L_z = \mathrm{dL}/\mathrm{d}z$, $\mathrm{G}_z = \mathrm{dG}/\mathrm{d}z$ and well known relation $\mathrm{dG}^{-1}/\mathrm{d}z = -\mathrm{G}^{-1}\mathrm{G}_z\mathrm{G}^{-1}$ were used. Now, using the relations $M = G_z G^{-1}$ and commutator bracket notation

$[M, L] = ML - LM$, it can be stated

$$\begin{aligned} \mathrm{L}_z &= \mathrm{G}_z \Lambda \mathrm{G}^{-1} + \mathrm{G}\Lambda(-\mathrm{G}^{-1}\mathrm{G}_z\mathrm{G}^{-1}) \\ &= \mathrm{M}\underbrace{(\mathrm{G}\Lambda\mathrm{G}^{-1})}_{\mathrm{L}} - \underbrace{(\mathrm{G}\Lambda\mathrm{G}^{-1})}_{\mathrm{L}}\underbrace{\mathrm{G}_z\mathrm{G}^{-1}}_{\mathrm{M}} \\ &= \mathrm{ML} - \mathrm{LM} = [M, L]. \end{aligned}$$

Similar derivations can be found in many related publications (e.g. in [18], [43]).

A.2. The Zero Curvature Condition

Starting from Eqs. (4.10) and (4.6) and again omitting temporal, spatial and λ dependencies in the notation, a set of mixed derivatives of v can be obtained in the form of

$$v_{t,z} = \mathrm{P}_z v + \mathrm{P} v_z, \tag{A.1}$$
$$v_{z,t} = \mathrm{M}_t v + \mathrm{M} v_t. \tag{A.2}$$

Replacing the derivatives of v on the right hand side, using Eqs. (4.10) and (4.6), results in

$$v_{t,z} = \mathrm{P}_z v + \mathrm{PM} v, \tag{A.3}$$
$$v_{z,t} = \mathrm{M}_t v + \mathrm{MP} v. \tag{A.4}$$

Equating the mixed derivatives stated above then yields the zero curvature condition. The operators $\mathrm{L}(z)$ and $\mathrm{P}(z)$ from Eqs. (4.5) and (4.10) can be related by reordering both equations to form

$$(\mathrm{L} - \lambda \mathrm{I})v = 0, \tag{A.5}$$
$$(\mathrm{DI} - \mathrm{P})v = 0. \tag{A.6}$$

Equating the above expressions after multiplication by some invertible operators Σ_1 and Σ_2 yields the relations

$$\begin{aligned}\Sigma_1(\mathrm{L}-\lambda\mathrm{I}) &= \Sigma_2(\mathrm{DI}-\mathrm{P}),\\ -\Sigma_2^{-1}\Sigma_1(\mathrm{L}-\lambda\mathrm{I})+\mathrm{DI} &= \mathrm{P},\\ \mathrm{P} &= \Sigma(\mathrm{L}-\lambda\mathrm{I})+\mathrm{DI},\end{aligned} \tag{A.7}$$

where $\Sigma = -\Sigma_2^{-1}\Sigma_1$ and Eq. (A.7) is exactly Eq. (4.12).

A.3. Properties of the Fundamental Soliton

In section 4.3.2, several quantities that can be computed in closed form for the fundamental soliton are introduced. In the following, the derivations or those quantities are given, using the notation from Eq. (4.62).

A.3.1. Temporal Width

Using Eq. (4.62) and (4.49), the relation (4.63) becomes

$$\begin{aligned}4\eta_1 p = &\int_{-T_p/2+t_0}^{T_p/2+t_0} \left|-\mathrm{j}\mathrm{e}^{-\mathrm{j}\phi_1}\mathrm{e}^{-\mathrm{j}4(\xi_1^2-\eta_1^2)z}\mathrm{e}^{-\mathrm{j}2\xi_1 t}2\eta_1\mathrm{sech}\left(2\eta_1 t+8\xi_1\eta_1 z-\ln\left(\frac{\hat{Q}_\mathrm{d}}{2\eta_1}\right)\right)\right|^2 \mathrm{d}t\\ =&\int_{-T_p/2}^{T_p/2} \left|2\eta_1\mathrm{sech}\left(2\eta_1 t'\right)\right|^2 \mathrm{d}t'\\ =&2\eta_1\left[\tanh\left(2\eta_1 t'\right)\right]_{-T_p/2}^{T_p/2} = 4\eta_1\tanh\left(2\eta_1 T/2\right),\end{aligned} \tag{A.8}$$

where in the second line a substitution was introduced to remove the influence of t_0. This can also be justified by just setting $t_0 = 0$, since a temporal shift of the pulse does not change the temporal width of the pulse itself. The last line above results in 4.64.

A.3.2. Bandwidth

Similarly, for the relation giving the linear bandwidth of a fundamental soliton, starting from Eq. (4.65) and the relation $\mathcal{F}\{\mathrm{sech}(ct+T)\} = \pi/c\cdot\mathrm{sech}(\pi^2/c\cdot f)\mathrm{e}^{-2\pi\mathrm{j}Tf}$, it can

be stated

$$\begin{aligned} 4\eta_1 p =& \int_{-B_p/2}^{B_p/2} \left| 2\eta_1 \frac{\pi}{2\eta_1} \operatorname{sech}\left(\frac{\pi^2}{2\eta_1} f\right) \mathrm{e}^{-2\pi \mathrm{j} T f} \right|^2 \mathrm{d}f \\ =& 2\eta \left[\tanh\left(\frac{\pi^2}{2\eta_1} f\right) \right]_{-B_p/2}^{B_p/2} = 4\eta_1 \tanh\left(\frac{\pi^2}{2\eta_1} \frac{B}{2}\right). \end{aligned} \tag{A.9}$$

From the last line the result given in Eq. (4.66) can be obtained.

A.4. 'Gray' Labeling

In section 4.11.1 it is stated, that by using the presented 'Gray'-labeling scheme, the bit-error per symbol-error can be reduced to one for the errors occuring with the highest probability according to the channel model defined in the beginning of section 4.11. In the following, a proof for this behavior is given and it is shown that generator matrix $\mathbf{G}_{\text{gray}}$ has the form presented in Eqs. (4.192) and (4.193).

First, to show that the transitions with highest error probabilities result in exactly one bit-error, Eq. (4.191) is restated as

$$\mathbf{u} = \mathbf{b} \cdot \mathbf{G}_{\text{gray}}^{-1}, \tag{A.10}$$

using the inverse generator matrix

$$\mathbf{G}_{\text{gray}}^{-1} = \begin{bmatrix} \mathbf{e}_N \\ \mathbf{e}_{N-1} \\ \vdots \\ \mathbf{e}_1 \end{bmatrix}, \tag{A.11}$$

where

$$\mathbf{e}_i = [\mathbf{0}_{N-i} \mathbf{1}_i]. \tag{A.12}$$

Parameters $\mathbf{0}_i$ and $\mathbf{1}_i$ denote all-zero and all-one row vectors of length i respectively. Rewriting Eq. (A.10) in the form

$$\mathbf{u} = [u_N u_{N-1} \cdots u_2 u_1] = b_N \mathbf{e}_N \oplus b_{N-1} \mathbf{e}_{N-1} \oplus \cdots \oplus b_2 \mathbf{e}_2 \oplus b_1 \mathbf{e}_1, \tag{A.13}$$

with $\oplus$ being the entry-wise XOR-operator. Using this representation all error types of the channel model with the highest probabilities can be shown to only result in one bit-error:

1. b_1 is flipped due to the discrete eigenvalue with the smallest imaginary part either vanishing or appearing during transmission. As a result, all terms in Eq. (A.13) remain unchanged except for the last one. Since $\mathbf{e}_1$ only has one non-zero position, altering s_1 only alters u_1.

2. $[b_i b_{i-1}] = [10] \rightarrow [b'_i b'_{i-1}] = [01]$ for $i \in \{2, 3, \ldots, N\}$, which represents the transmitted discrete eigenvalue vanishing and the neighboring lower energy discrete eigenvalue appearing. In this case term $b_i \mathbf{e}_i$ is replaced by $b_{i-1}\mathbf{e}_{i-1}$. Vectors $\mathbf{e}_i$ and $\mathbf{e}_{i-1}$ only differ in position i and thus only u_i is altered in this case.

3. $[b_{i+1} b_i] = [01] \rightarrow [b'_{i+1} b'_i] = [10]$ for $i \in \{1, 2, \ldots, N-1\}$, which represents the transmitted discrete eigenvalue vanishing and the neighboring higher energy discrete eigenvalue appearing. Similar to the previous case, term $b_{i+1}\mathbf{e}_{i+1}$ replaces $b_i \mathbf{e}_i$. Vectors $\mathbf{e}_{i+1}$ and $\mathbf{e}_i$ only differ in position $i+1$ and thus only u_{i+1} is altered.

In the next step, it remains to be shown that $\mathbf{G}_{\text{gray}}$ is given by Eqs. (4.192) and (4.193). Starting with

$$\mathbf{G}_{\text{gray}} \mathbf{G}_{\text{gray}}^{-1} = \mathbf{I}_N, \tag{A.14}$$

it can be seen that this results in the condition

$$\mathbf{e}_i \mathbf{d}_j = \begin{cases} 1, & \text{for } i = j \\ 0, & \text{otherwise} \end{cases}, \tag{A.15}$$

where $\mathbf{d}_j$ are column vectors that will be shown to have the form given in Eq. (4.193). First it is assumed that $j = N$. For this case, it can be written

$$\mathbf{e}_i \cdot \mathbf{d}_j = [\mathbf{0}_{N-i} \mathbf{1}_i] \cdot \begin{bmatrix} 1 \\ \mathbf{0}_{N-1}^{\mathrm{T}} \end{bmatrix} = \begin{cases} 1, & \text{for } i = j = N \\ 0, & \text{otherwise} \end{cases}. \tag{A.16}$$

It can be seen, that for the case $i = j = N$ vector $\mathbf{e}_N = \mathbf{1}_N$ and thus $\mathbf{e}_N \cdot \mathbf{d}_N = 1$. In all other cases the leftmost entry of $\mathbf{e}_i$ is zero and thus $\mathbf{e}_i \cdot \mathbf{d}_N = 0$ for $i \neq N$.

For the remaining cases $j = 1, \cdots, N-1$ Eq. (A.15) can be rewritten as

$$\mathbf{e}_i \cdot \mathbf{d}_j = [\mathbf{0}_{N-i} \mathbf{1}_i] \cdot \begin{bmatrix} \mathbf{0}_{N-j-1}^{\mathrm{T}} \\ 1 \\ 1 \\ \mathbf{0}_{j-1}^{\mathrm{T}} \end{bmatrix} = \begin{cases} 1, & \text{for } i = j \\ 0, & \text{otherwise} \end{cases} . \tag{A.17}$$

In $\mathbf{d}_j$, all entries are zero except for the entries in positions j and $j+1$ being one. For vector $\mathbf{e}_i$ the i rightmost entries are one and the rest is zero. Thus, if $i < j$, both non-zero entries in $\mathbf{d}_j$ are multiplied with zeros and as a result $\mathbf{e}_i \cdot \mathbf{d}_j = 0$. Similarly, for $i > j$ both non-zero values in $\mathbf{d}_j$ are multiplied with ones. Thus $\mathbf{e}_i \cdot \mathbf{d}_j = 0$ is still true in these cases. For the remaining case of $i = j$, one of the non-zero entries in $\mathbf{d}_j$ is cancelled out by zeros in $\mathbf{e}_i$ and thus $\mathbf{e}_i \cdot \mathbf{d}_j = 1$. Since Eqs. (A.16) and (A.17) are fulfilled for the respective vectors used, it has been shown that generator matrix $\mathbf{G}_{\text{gray}}$ is given by Eqs. (4.192) and (4.193).

A.5. Properties of Nonlinear Spectra in the Multi-Mode Case

Several useful properties for the scalar NFT exist (as shown in e.g. [43]). A similar set of properties for the multi-mode case was presented in section 5.2. In the following, the proofs for the respective properties are given, using the notation introduced in section 5.2. Note, that the notation used will be simplified by e.g. omitting unnecessary indexing or dependencies for the sake of readability.

A.5.1. Differential Representation for Multi-Mode NFT

The differential representation for the continuous spectrum, analog to [43], can be obtained by starting at $\boldsymbol{v_t} = \mathbf{P}\boldsymbol{v}$ with $\boldsymbol{v}$ being a $M+1$ column vector with entries v_i ($0 \leq i \leq M$) and $\mathbf{P}$ defined according to 5.44. The evaluation of each row yields

$$v_{i,t} = \begin{cases} -\mathrm{j}\lambda v_i + \sum\limits_{m=1}^{M} q_m v_m & \text{for } i = 0 \\ -q_i^* v_0 + \mathrm{j}\lambda v_i & \text{for } 1 \leq i \leq M \end{cases} \tag{A.18}$$

Now the quantity

$$y_m = \frac{v_m}{v_0} \mathrm{e}^{-2\mathrm{j}\lambda t}, \tag{A.19}$$

is defined. It can be seen that $q_{\mathrm{c}}^{(m)}(\lambda) = \lim_{t \to +\infty} y_m(t, \lambda)$. In the next step, the derivative

$$\frac{\mathrm{d}y_m}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{v_m}{v_0} \mathrm{e}^{-2\mathrm{j}\lambda t} \right) \tag{A.20}$$

$$= \frac{(v_{m,t} \mathrm{e}^{-2\mathrm{j}\lambda t} + v_m(-2\mathrm{j}\lambda)\mathrm{e}^{-2\mathrm{j}\lambda t}) v_0 - v_m \mathrm{e}^{-2\mathrm{j}\lambda t} v_{0,t}}{v_0^2}, \tag{A.21}$$

can be computed. Substituting the temporal derivatives of v_i using Eq. (A.18) and then expressing all v_i parameters as y_m using Eq. (A.19) yields the differential form

$$\frac{\mathrm{d}y_m}{\mathrm{d}t} + y_m \cdot \sum_{i=1}^{M} q_i y_i \mathrm{e}^{2\mathrm{j}\lambda t} + q_m^* \mathrm{e}^{-2\mathrm{j}\lambda t} = 0. \tag{A.22}$$

The boundary condition, given by Eq. (5.13), can be verified directly, by using Eq. (5.8) in Eq. (A.22).

Similarly, a second order differential equation for the multi-mode case can be found, yielding the parameter $a(\lambda)$. This, again, is the analog to the second order differential equation derived for the single-mode case in [43]. For the sake of brevity, the first and second order temporal derivatives are notated by $\mathrm{d}x/\mathrm{d}t = x'$ and $\mathrm{d}^2x/\mathrm{d}t^2 = x''$ respectively. Defining $z = v_0 \mathrm{e}^{\mathrm{j}\lambda t}$, the first and second order derivatives are

$$z' = v_0' \mathrm{e}^{\mathrm{j}\lambda t} + v_0 \mathrm{j}\lambda \mathrm{e}^{\mathrm{j}\lambda t}, \tag{A.23}$$

$$z'' = v_0'' \mathrm{e}^{\mathrm{j}\lambda t} + v_0' 2\mathrm{j}\lambda \mathrm{e}^{\mathrm{j}\lambda t} + v_0 (\mathrm{j}\lambda)^2 \mathrm{e}^{\mathrm{j}\lambda t}. \tag{A.24}$$

The first order derivatives of v_i are given in Eq. (A.18). The missing second order derivative of the first entry of the eigenvector is computed to be

$$v_0'' = -\mathrm{j}\lambda v_0' + \sum_{m=1}^{M} q_m' v_m + \sum_{m=1}^{M} q_m v_m'. \tag{A.25}$$

Inserting Eq. (A.25) into Eq. (A.24) yields

$$\begin{aligned} z'' &= (-\mathrm{j}\lambda v_0' + \sum_{m=1}^{M} q_m' v_m + \sum_{m=1}^{M} q_m v_m')\mathrm{e}^{\mathrm{j}\lambda t} + v_0'\mathrm{j}\lambda\mathrm{e}^{\mathrm{j}\lambda t} + \underbrace{v_0'\mathrm{j}\lambda\mathrm{e}^{\mathrm{j}\lambda t} + v_0(\mathrm{j}\lambda)^2\mathrm{e}^{\mathrm{j}\lambda t}}_{\mathrm{j}\lambda z'} \\ &= (\sum_{m=1}^{M} q_m' v_m + \sum_{m=1}^{M} q_m v_m')\mathrm{e}^{\mathrm{j}\lambda t} + \mathrm{j}\lambda z'. \end{aligned} \tag{A.26}$$

Further, inserting v_0' from Eq. (A.18) results in

$$\begin{aligned} z'' - \mathrm{j}\lambda z' &= (\sum_{m=1}^{M} q_m' v_m + \sum_{m=1}^{M} q_m(-q_m^* v_0 + \mathrm{j}\lambda v_m))\mathrm{e}^{\mathrm{j}\lambda t} \\ &= \sum_{m=1}^{M} q_m' v_m \mathrm{e}^{\mathrm{j}\lambda t} - \sum_{m=1}^{M} |q_m|^2 \underbrace{v_0 \mathrm{e}^{\mathrm{j}\lambda t}}_{z} + \underbrace{\sum_{m=1}^{M} \mathrm{j}\lambda q_m v_m \mathrm{e}^{\mathrm{j}\lambda t}}_{\mathrm{j}\lambda z'}. \end{aligned}$$

Using the definitions for z and z' the equation above can be rearranged, stating

$$z'' - 2\mathrm{j}\lambda z' + \sum_{m=1}^{M} |q_m|^2 z - \sum_{m=1}^{M} q_m' v_m \mathrm{e}^{\mathrm{j}\lambda t} = 0. \tag{A.27}$$

This result is less compact than the result in [43] and not all parameters v_i can be absorbed in the defined parameters z. This fact potentially makes the multi-mode version of the second order differential representation less useful. Nonetheless, the result in Eq. (A.27) can be modified, such that it is more similar to the result in [43]. By expanding the last term in Eq. (A.27) as

$$\begin{aligned} \sum_{m=1}^{M} q_m' v_m \mathrm{e}^{\mathrm{j}\lambda t} &= \sum_{m=1}^{M} \frac{q_m'}{q_m} \left(v_0' + \mathrm{j}\lambda v_0 - \sum_{\substack{n=1 \\ n\neq m}}^{M} q_n v_n \right) \mathrm{e}^{\mathrm{j}\lambda t} \\ &= \sum_{m=1}^{M} \frac{q_m'}{q_m} z' - \sum_{m=1}^{M} \sum_{\substack{n=1 \\ n\neq m}}^{M} \frac{q_m'}{q_m} q_n v_n \mathrm{e}^{\mathrm{j}\lambda t}, \end{aligned}$$

Eq. (A.27) can be restated as

$$z'' - \left(2\mathrm{j}\lambda + \sum_{m=1}^{M} \frac{q_m'}{q_m} \right) z' + \sum_{m=1}^{M} |q_m|^2 z + \sum_{m=1}^{M} \sum_{\substack{n=1 \\ n\neq m}}^{M} \frac{q_m'}{q_m} q_n v_n \mathrm{e}^{\mathrm{j}\lambda t} = 0. \tag{A.28}$$

It can be seen, that in the case of $M = 1$ this is identical to the result from [43]. The boundary conditions

$$z(t \to -\infty, \lambda) = 1, \tag{A.29}$$

$$z'(t \to -\infty, \lambda) = 0, \tag{A.30}$$

can be verified easily by looking at the definition of z, Eq. (A.23) and the boundary condition $v_0(t \to -\infty, \lambda) = \mathrm{e}^{-\mathrm{j}\lambda t}$.

A.5.2. Constant Phase Change

The constant phase change property, derived for the single-mode case in [43], can be extended to the multi-mode case. This can be shown, by replacing $q_m(t) \to q_m(t)\mathrm{e}^{\mathrm{j}\phi_m}$ in Eq. (A.22) and rearranging the equation until $y_m \to y_m \mathrm{e}^{\mathrm{j}\phi}$. This implies that $q_m(t) \to q_m(t)\mathrm{e}^{\mathrm{j}\phi_m}$ leads to $q_\mathrm{c}^{(m)}(\lambda) \to q_\mathrm{c}^{(m)}(\lambda)\mathrm{e}^{-\mathrm{j}\phi_m}$. Another way of showing that this property holds in the multi-mode case, is to modify the temporal propagation equation for eigenvectors. By doing so, a proof of a similar property for the discrete spectrum parameters is obtained as well. Exchanging $q_m(t) \to q_m(t)\mathrm{e}^{\mathrm{j}\phi_m}$ in the P operator and defining

$$u = \mathrm{diag}(1, \mathrm{e}^{\mathrm{j}\phi_1}, \mathrm{e}^{\mathrm{j}\phi_2}, \cdots, \mathrm{e}^{\mathrm{j}\phi_M})v, \tag{A.31}$$

it can be shown that u fulfills the temporal propagation equation for the modified P operator. Since, u and v are related by Eq. (A.31), it follows that the nonlinear Fourier coefficients are modified according to $b_m \to b_m \mathrm{e}^{-\mathrm{j}\phi_m}$.

A.5.3. Time Shift

The time shift property for the multi-mode case can be proven by representing the nonlinear Fourier coefficients by the propagator $\boldsymbol{K}$ and the initial values of the nonlinear Fourier coefficients. It can be stated that

$$\begin{bmatrix} a \cdot \mathrm{e}^{-\mathrm{j}\lambda t_2} & b_1 \cdot \mathrm{e}^{\mathrm{j}\lambda t_2} & b_2 \cdot \mathrm{e}^{\mathrm{j}\lambda t_2} & \cdots & b_M \cdot \mathrm{e}^{\mathrm{j}\lambda t_2} \end{bmatrix}^\mathrm{T} = \boldsymbol{K} \begin{bmatrix} \mathrm{e}^{-\mathrm{j}\lambda t_1} & 0 & 0 & \cdots & 0 \end{bmatrix}^\mathrm{T}, \tag{A.32}$$

where

$$\boldsymbol{K} = \begin{bmatrix} a \cdot \mathrm{e}^{-\mathrm{j}\lambda(t_2-t_1)} & -b_1^* \cdot \mathrm{e}^{-\mathrm{j}\lambda(t_2+t_1)} & -b_2^* \cdot \mathrm{e}^{-\mathrm{j}\lambda(t_2+t_1)} & \cdots & -b_M^* \cdot \mathrm{e}^{-\mathrm{j}\lambda(t_2+t_1)} \\ b_1 \cdot \mathrm{e}^{\mathrm{j}\lambda(t_2+t_1)} & a^* \mathrm{e}^{\mathrm{j}\lambda(t_2-t_1)} & 0 & \cdots & 0 \\ b_2 \cdot \mathrm{e}^{\mathrm{j}\lambda(t_2+t_1)} & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ b_M \cdot \mathrm{e}^{\mathrm{j}\lambda(t_2+t_1)} & 0 & \cdots & 0 & a^* \mathrm{e}^{\mathrm{j}\lambda(t_2-t_1)} \end{bmatrix}. \tag{A.33}$$

If it is now assumed that the nonlinear Fourier coefficients in Eq. (A.32) and (A.33) belong to pulse-set $\boldsymbol{q}(t)$ and pulse-set $\hat{\boldsymbol{q}}(t) = \boldsymbol{q}(t-t_0)$ has nonlinear Fourier coefficients $\hat{a}(\lambda)$, $\hat{b}_m(\lambda)$. Further, the time variables, regarding the original and the shifted pulse-set, are related by $\hat{t}_1 = t_1 + t_0$, $\hat{t}_2 = t_2 + t_0$. To arrive at a system similar to Eqs. (A.32) and (A.33) for the shifted pulse, it can be seen that relations

$$\hat{a}(\lambda) = a(\lambda)\mathrm{e}^{\mathrm{j}\lambda t_0}, \tag{A.34}$$

$$\hat{b}_m(\lambda) = b_m(\lambda)\mathrm{e}^{-\mathrm{j}\lambda t_0}, \tag{A.35}$$

have to be fulfilled. Using Eq. (5.11) the validity of Eq. (5.23) can be shown as well.

A.5.4. Trace Formula

In the following, the trace formula is used to derive the energy relation between the time-domain pulse-set $\boldsymbol{q}(t)$ and the continuous and discrete spectrum components $q_{\mathrm{c}}^{(m)}(\lambda)$, $q_{\mathrm{d}}^{(m)}(\lambda_k)$ defined according to Eq. (5.11). The starting point is the general trace formula given by [90, Ch. 1.6], evaluated for $n = 0$

$$E_{\mathrm{c}} = E(\boldsymbol{q}_{\mathrm{c}}(\lambda)) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \log(|a(\lambda)|^2)\mathrm{d}\lambda, \tag{A.36}$$

$$E_{\mathrm{d}} = E(\boldsymbol{q}_{\mathrm{d}}(\lambda_k)) = \sum_{k=1}^{K} \left[(2\mathrm{j}\lambda_k^*) - (2\mathrm{j}\lambda_k)\right]. \tag{A.37}$$

For Eq. (A.37) it can easily be seen, that it evaluates to the well known expression

$$E_{\mathrm{d}} = 4\sum_{k=1}^{K} \Im\{\lambda_k\}, \tag{A.38}$$

which is identical for the single and multi-mode NFT. The continuous spectrum energy relation can be obtained by first using the unimodularity condition given in 5.2.3 to state

$$1 + \sum_{m=1}^{M} |q_{\mathrm{c}}^{(m)}(\lambda)|^2 = \frac{1}{|a(\lambda)|^2}. \tag{A.39}$$

Using the logarithm on the above equation, yields

$$\log\left(1 + \sum_{m=1}^{M} |q_{\mathrm{c}}^{(m)}(\lambda)|^2\right) = -\log\left(|a(\lambda)|^2\right). \tag{A.40}$$

It it easy to see, that using Eq. (A.40) in Eq. (A.36) results in the energy relation for the continuous nonlinear spectrum

$$E_{\mathrm{c}} = \frac{1}{\pi} \int_{-\infty}^{\infty} \log\left(1 + \sum_{m=1}^{M} |q_{\mathrm{c}}^{(m)}(\lambda)|^2\right) \mathrm{d}\lambda. \tag{A.41}$$

A.6. NFT of the Vectorial Rectangular Pulse

Starting from

$$\boldsymbol{v_t}(t, \lambda) = \mathbf{P}(z)\boldsymbol{v}(t, \lambda) \tag{A.42}$$

with operator $\mathbf{P}(z)$ defined by Eq. (5.44), it can be seen that for a rectangular pulse-set, the operator is constant over the pulse-set duration for constant z. Note, that this only is true, if the rectangular pulses in the set have equal temporal support as defined in section 5.3.1. As a result, the values of $\boldsymbol{v}(t, \lambda)$ for the temoral positions t_1, t_2 can be linked by

$$\boldsymbol{v}(\lambda, t_2) = \boldsymbol{v}(\lambda, t_1) \cdot \mathrm{e}^{\boldsymbol{P}T}. \tag{A.43}$$

The matrix exponential can be evaluated using the Taylor-series representation

$$\mathrm{e}^{\boldsymbol{P}T} = \sum_{i=0}^{\infty} \frac{T^i}{i!} \boldsymbol{P}^i, \tag{A.44}$$

where $x!$ is the faculty of x. It can be seen, that the powers of operator $\boldsymbol{P}$ are needed. While tedious, these can be computed fairly straightforward. The entries for the first few powers of $\boldsymbol{P}$ are given explicitely in the following, denoting the entries of $\boldsymbol{P}^i$ by $(\boldsymbol{P}^i)_{kl}$. For the sake of ease of notation, the entries will be labeled starting with column and row index zero, as will be clearly denoted in the following. For the power of two

the entries are

$$(\boldsymbol{P}^2)_{00} = -\Delta_\Sigma, \tag{A.45}$$
$$(\boldsymbol{P}^2)_{kk} = -\Delta_k \qquad \text{for } k = 1, \cdots, M, \tag{A.46}$$
$$(\boldsymbol{P}^2)_{0k} = 0 \qquad \text{for } k = 1, \cdots, M, \tag{A.47}$$
$$(\boldsymbol{P}^2)_{k0} = 0 \qquad \text{for } k = 1, \cdots, M, \tag{A.48}$$
$$(\boldsymbol{P}^2)_{kl} = -A_k^* A_l \qquad \text{for } k, l = 1, \cdots, M,\ k \neq l. \tag{A.49}$$

For $\boldsymbol{P}^3$ the entries are

$$(\boldsymbol{P}^3)_{00} = \mathrm{j}\lambda\Delta_\Sigma, \tag{A.50}$$
$$(\boldsymbol{P}^3)_{kk} = -\mathrm{j}\lambda\Delta_1 \qquad \text{for } k = 1, \cdots, M, \tag{A.51}$$
$$(\boldsymbol{P}^3)_{0k} = -A_k\Delta_\Sigma \qquad \text{for } k = 1, \cdots, M, \tag{A.52}$$
$$(\boldsymbol{P}^3)_{k0} = A_k^*\Delta_\Sigma \qquad \text{for } k = 1, \cdots, M, \tag{A.53}$$
$$(\boldsymbol{P}^3)_{kl} = -\mathrm{j}\lambda A_k^* A_l \qquad \text{for } k, l = 1, \cdots, M,\ k \neq l. \tag{A.54}$$

For $\boldsymbol{P}^4$ the entries are

$$(\boldsymbol{P}^4)_{00} = \Delta_\Sigma^2, \tag{A.55}$$
$$(\boldsymbol{P}^4)_{kk} = \Delta_k^2 + |A_k|^2\Sigma_{m\neq k}^M \qquad \text{for } k = 1, \cdots, M, \tag{A.56}$$
$$(\boldsymbol{P}^4)_{0k} = 0 \qquad \text{for } k = 1, \cdots, M, \tag{A.57}$$
$$(\boldsymbol{P}^4)_{k0} = 0 \qquad \text{for } k = 1, \cdots, M, \tag{A.58}$$
$$(\boldsymbol{P}^4)_{kl} = (2\lambda^2 + \Sigma)A_k^* A_l \qquad \text{for } k, l = 1, \cdots, M,\ k \neq l. \tag{A.59}$$

The newly introduced parameters in the above equations are defined as

$$\Sigma = \sum_{m=1}^{M} |A_m|^2, \tag{A.60}$$
$$\Delta_m = \lambda^2 + |A_m|^2, \tag{A.61}$$
$$\Delta_\Sigma = \lambda^2 + \Sigma, \tag{A.62}$$
$$\Sigma_{m\neq i}^M = -|A_i|^2 + \sum_{m=1}^{M} |A_m|^2. \tag{A.63}$$

Using Eqs. (A.45) - (A.59) and noting that the Taylor-series expansions of sine and cosine functions are

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots, \tag{A.64}$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots, \tag{A.65}$$

the first column of the matrix-exponential can be computed. The first column specifically is chosen, since it corresponds with the scaled nonlinear Fourier coefficients $a(\lambda)\mathrm{e}^{\mathrm{j}\lambda(t_2-t_1)}$ and $b_m(\lambda)\mathrm{e}^{-\mathrm{j}\lambda(t_2+t-1)}$. Rearranging the obtained results, yields the closed form equations for the nonlinear Fourier coefficients of the rectangular pulse-set

$$a(\lambda) = (\cos(\Delta T) - \mathrm{j}\frac{\lambda}{\Delta}\sin(\Delta T))\mathrm{e}^{\mathrm{j}\lambda(t_2-t_1)} \tag{A.66}$$

$$b_m(\lambda) = -\frac{A_m^*}{\Delta}\sin(\Delta T)\mathrm{e}^{-\mathrm{j}\lambda(t_2+t_1)}, \tag{A.67}$$

where $\Delta = \sqrt{\lambda^2 + \sum_{m=1}^{M} |A_m|^2}$. Further, Eq. (5.37), which can be used to find the zeros $a(\lambda)$ is obtained by setting Eq. (A.66) to zero.

Note, that it is also possible to give a closed form expression for $a_\lambda(\lambda_k)$, which is the λ-derivative of Eq. (A.66) evaluated for discrete eigenvalue λ_k. This expression then can be used with Eq. (A.67) to obtain the discrete spectral amplitudes $q_d^{(m)}(\lambda_k) = b_m(\lambda_k)/a_\lambda(\lambda_k)$. Since these specific parameters are not of significant interest for this thesis, any further treatment of them is omitted.

A.7. Properties of the Vectorial Fundamental Soliton

In the following, the derivations for the properties of the vectorial fundamental soliton as treated in Section 5.3.2 are given, using the notation introduced in Eqs. (5.38) and (5.39).

A.7.1. Temporal Width

Starting from Eq. (5.39) and inserting (5.38), the propagation distance z is set to zero, since it does not influence the temporal width if the temporal width is defined as the width necessary to support factor p of the pulses energy, assuming the pulse is

centered on the temporal support. Obviously, this only is true for z-propagation in the ideal channel. After dropping the phase term inside the absolute value operator and defining $\Delta = \sum_{s=1}^{M} |Q_d^{(s)}/2\eta|^2$

$$E_{\text{vfs}} \cdot p = 4\eta^2 \int_{-T/2+c}^{T/2+c} \sum_{m=1}^{M} \left| \frac{2\mathrm{e}^{-2\eta t}}{1+\mathrm{e}^{-4\eta t}\Delta} \right|^2 \left| \frac{Q_d^{(m)}}{2\eta} \right|^2 \mathrm{d}t, \tag{A.68}$$

$$= 4\eta^2 \int_{-T/2+c}^{T/2+c} \left| \frac{2\mathrm{e}^{-2\eta t}\sqrt{\Delta}}{1+\mathrm{e}^{-4\eta t}\Delta} \right|^2 \mathrm{d}t, \tag{A.69}$$

$$= 4\eta^2 \int_{-T/2+c}^{T/2+c} \left| \operatorname{sech}(2\eta t - \frac{1}{2}\ln(\Delta)) \right|^2 \mathrm{d}t, \tag{A.70}$$

can be obtained. It can be seen, that $c = 0.5 \cdot \ln(\Delta)/2\eta$ is the pulse center and thus

$$E_{\text{vfs}} \cdot p = \left[\frac{\tanh(2\eta t)}{2\eta} \right]_{-T/2}^{T/2} \tag{A.71}$$

$$= 4\eta \cdot \tanh(\eta T). \tag{A.72}$$

With $E_{\text{vfs}} = E_{\text{d}} = 4\eta$ according to Eq. (5.25) and the fact that the fundamental soliton set only has one discrete eigenvalue $\lambda_1 = \mathrm{j}\eta$, the temporal width according to some in-window energy restriction p is given by

$$T = \frac{\operatorname{arctanh}(\mathrm{p})}{\eta}. \tag{A.73}$$

A.7.2. Bandwidth

Inserting Eq. (5.38) into (5.41), setting $z = 0$ and dropping time-independent terms, that are equal to one, due to the absolute value operation if pulled out of the Fourier transform it can be stated that

$$E_{\text{vfs}} \cdot p = 4\eta^2 \int_{-B/2+c}^{B/2+c} \sum_{m=1}^{M} \left| \mathcal{F}\left\{ \mathrm{e}^{-\mathrm{j}2\xi t} \cdot \frac{2\mathrm{e}^{-2\eta t}\Delta_m}{1+\mathrm{e}^{-4\eta t}\Delta} \right\} \right|^2 \mathrm{d}f, \tag{A.74}$$

where $\Delta_m = |Q_d^{(m)}/2\eta|^2$ and Δ is defined as in A.7.1. It can be seen, that the ξ-dependent exponent is causing a frequency shift ξ/π in the Fourier transform. The remaining term is transformed according to

$$\begin{aligned}\mathcal{F}\left\{\frac{2\mathrm{e}^{-2\eta t}\Delta_m}{1+\mathrm{e}^{-4\eta t}\Delta}\right\} =& \frac{\Delta_m}{\sqrt{\Delta}}\cdot\mathcal{F}\left\{\frac{2\mathrm{e}^{-2\eta t}\sqrt{\Delta}}{1+\mathrm{e}^{-4\eta t+\ln(\Delta)}}\right\}\\ =& \frac{\Delta_m}{\sqrt{\Delta}}\cdot\mathcal{F}\left\{\operatorname{sech}(2\eta t-\frac{1}{2}\ln(\Delta))\right\}\\ =& \frac{\Delta_m}{\sqrt{\Delta}}\cdot\left(\frac{\pi}{2\eta}\cdot\operatorname{sech}\left(\frac{\pi^2}{2\eta}f\right)\right)\cdot\mathrm{e}^{-\mathrm{j}2\pi\ln(\Delta)f}\end{aligned} \tag{A.75}$$

Inserting Eq. (A.75) in (A.74) and removing the phase term due to the $|\cdot|$ operation, the energy bandwidth relation is computed according to

$$\begin{aligned}E_{\mathrm{vfs}}\cdot p =& 4\eta^2\int_{-B/2+c}^{B/2+c}\sum_{m=1}^{M}\left|\frac{\Delta_m}{\Delta}\left(\frac{\pi}{2\eta}\operatorname{sech}\left(\frac{\pi^2}{2\eta}\left(f+\frac{\xi}{\pi}\right)\right)\right)\right|^2\mathrm{d}f\\ =& \pi^2\frac{\sum_{m=1}^{M}|\Delta_m|^2}{\Delta}\int_{-B/2+c}^{B/2+c}\left|\operatorname{sech}\left(\frac{\pi^2}{2\eta}\left(f+\frac{\xi}{\pi}\right)\right)\right|^2\mathrm{d}f\\ =& \pi^2\left[\frac{\tanh\left(\frac{\pi^2}{2\eta}f\right)}{\frac{pi^2}{2\eta}}\right]_{-B/2}^{B/2} = 4\eta\tanh\left(\frac{\pi^2}{2\eta}\frac{B}{2}\right).\end{aligned}$$

Rearranging the last step in the equation above, once more using the energy relation $E_{\mathrm{vfs}} = E_{\mathrm{d}} = 4\eta$ according to Eq. (5.25), finally yields

$$B = \frac{4\eta}{\pi^2}\cdot\operatorname{arctanh}(p). \tag{A.76}$$

A.8. Algorithms for the Multi-Mode Case

Over the course of this thesis, only a selected set of algorithms was extended to be usable in the multi-mode case. Naming conventions, as well as notation was kept as similar as possible to the single-mode case. The derivation of the FD-method is very similar to the single-mode case, and so straightforward that we deem it unnecessary to reiterate it explicitly in this appendix. In the following, the necessary equations for the LP-method are derived. For the derivations, the notation from sections 5.4.1

and 5.4.2 are used. The notation is further simplified, where necessary, for the sake of better readability. If a temporal pulse-set $\boldsymbol{q}(t) \rightarrow \{a(\lambda), b_m(\lambda)\}$ is approximated by N non-overlapping adjacent rectangular pulse-sets $\boldsymbol{q}^{(n)}(t) \rightarrow \{x^{(n)}(\lambda), y_m^{(n)}(\lambda)\}$ for $n = 1, \cdots, N$, the LP-property from Eqs. (5.26) and (5.27) and the initial values for the nonlinear Fourier coefficients $a^{(0)}(\lambda) = 1$, $b_m^{(0)}(\lambda) = 0 \,\forall\, m = 1, \cdots, M$, can be used to construct an iterative algorithm for the approximation of the nonlinear Fourier coefficients of an arbitrary pulse-set $\boldsymbol{q}(t)$. The update Eqs. (5.26) and (5.27) can be restated as

$$a^{(n)} = a^{(n-1)} x^{(n)} - \sum_{m=1}^{M} b_m^{(n-1)} y_m^{(n),*}, \tag{A.77}$$

$$b_m^{(n)} = a^{(n-1)} y_m^{(n)} + b_m^{(n-1)} x^{(n),*}. \tag{A.78}$$

The nonlinear Fourier coefficients are then approximated according to $\{a(\lambda), b_m(\lambda)\} \approx \{a^{(N)}(\lambda), b_m^{(N)}(\lambda)\}$. The update coefficients for the iteratively added rectangular pulses can be obtained straightforward, using Eqs. (A.66) and (A.67). The resulting relations are given in Eqs. (5.51) and (5.52). In a next step, the relations needed or the discrete spectrum computation can be derived. This is achieved by the straightforward computation of λ-derivatives of the update Eqs. (A.77), (A.78), (5.51) and (5.52), resulting in Eqs. (5.57), (5.58), (5.59) and (5.60) respectively.

B

Additional Algorithms

B.1. Alternative Continuous Spectrum RH-Solver

In this section, an alternative version of the RH-solver from section 4.4.5 is presented. The notation is adapted from the aforementioned section and simplified for the sake of better readability. Note, that the benefits and drawbacks of either of the two versions is not studied in this thesis. This is the main reason, this alternative version is just briefly mentioned in the appendix. First, Eqs. (4.145) and (4.146) are used to define

$$\boldsymbol{J}_1 = \underbrace{[\boldsymbol{B}_1,\, \boldsymbol{B}_1 \cdots ,\, \boldsymbol{B}_1]}_{2\mathrm{x}N}, \tag{B.1}$$

$$\boldsymbol{J}_2 = \underbrace{[\boldsymbol{B}_2,\, \boldsymbol{B}_2 \cdots ,\, \boldsymbol{B}_2]}_{2\mathrm{x}N}. \tag{B.2}$$

Using Eq. (4.147), it can be stated that

$$\boldsymbol{\mathcal{V}}^1 = \boldsymbol{J}_2 + \tilde{\boldsymbol{\mathcal{V}}}^1 \boldsymbol{C}^*, \tag{B.3}$$

$$\tilde{\boldsymbol{\mathcal{V}}}^1 = \boldsymbol{J}_1 + \boldsymbol{\mathcal{V}}^1 \boldsymbol{C}. \tag{B.4}$$

If now Eq. (B.4) is inserted into Eq. (B.3), the relation

$$\boldsymbol{\mathcal{V}}^1 = (\boldsymbol{J}_2 + \boldsymbol{J}_1 \boldsymbol{C}^*)(\boldsymbol{I} - \boldsymbol{C}\boldsymbol{C}^*)^{-1}, \tag{B.5}$$

can be obtained. This result the can be used to either recover the nonlinear Fourier coefficients by using Eq. (4.144) or recover the time-domain signal via Eq. (4.148).

B.2. Greedy Labeling Algorithm

The greedy labeling algorithm used in this thesis, is finding the biggest transition probability in transition probability matrix $\mathbf{P}$ and assigns some optimum bit labeling to the corresponding symbols if possible. The algorithm used is given in algorithm 2 in pseudo-code. Parameters $\boldsymbol{s}_\mathrm{P}$ and $\boldsymbol{r}_\mathrm{P}$ are the transmitted and received pilot symbols respectively and M is the cardinality of the symbol alphabet. Note, that functions starting with the prefix "*pop*" also remove the returned labels from $\mathbf{g}$. Also, text after "//" signs is a comment and does not belong to the functional part of the shown pseudo-code.

Algorithm 2: Greedy labeling algorithm

```
Result: l // labeling vector
P = CalculateTransitionProbabilityMatrix(s_P, r_P);
g = GenerateAllPossibleBitLabels(M);
while not all symbols have been labeled do
    [i, j] = FindIndicesOfMaximumEntry(P);
    P(i, j) = 0 // mark as used for future iterations;
    if symbols i and j unlabeled then
        [l(i), l(j), g] =popRemainingLabelPairWithMinimumBitDifference(g);
    else if symbol i already labeled, symbol j unlabeled then
        [l(j), g] =popRemainingLabelWithMinimumBitDifference(l(i), g);
    else if symbol j already labeled, symbol i unlabeled then
        [l(i), g] =popRemainingLabelWithMinimumBitDifference(l(j), g);
    else
        // both symbols already labeled → do nothing
    end
    if all transition probabilities in P are zero then
        [l] = assignRemainingLabelsToUnlabeledSymbolsArbitrarily(P, g);
end
```

B.3. Marching Squares Algorithm

After the detection of phase jumps $\{g_k\}$ on ∂G, using algorithm 1 in the PJT scheme presented in section 4.4.2, the phase jump trajectories have to be tracked back to their origins. At these points of origin, the discrete eigenvalues λ_k of the discrete nonlinear Fourier spectrum are located. Generally, there are several suitable algorithms to traverse these trajectories. In this thesis, the relatively simple and well known *marching squares* algorithm [112] is used. In algorithm 3 a straightforward implementation is given, using the notation from section 4.4.2.

Algorithm 3: Phase Jump Traversal (adapted from [111])

Result: $\{\hat{\lambda}_k\}$
for $k = 1, \cdots, K$ **do**
  $l = g_k - 0.5h_\lambda e^{j\varphi_k}$; $r = g_k + 0.5h_\lambda e^{j\varphi_k}$;
  while true **do**
    $s = h_\lambda e^{j(\varphi_k+\pi/2)}$; $\tilde{r} = r + s$;
    if isJump$(r, \tilde{r})$ **then**
      $l = \tilde{r}$; $\varphi_k = \varphi_k - \pi/2$;
    else
      $\tilde{l} = l + s$;
      if isJump$(\tilde{l}, \tilde{r})$ **then**
        $l = \tilde{l}$; $r = \tilde{r}$;
      else
        if isJump$(l, \tilde{l})$ **then**
          $r = \tilde{l}$; $\varphi_k = \varphi_k + \pi/2$;
        else
          $\lambda_k = 0.25 \cdot (l + r + \tilde{l} + \tilde{r})$; break;
        end
      end
    end
  end
end

In the algorithm above h_λ is the step size, which is defined as

$$h_\lambda = C_h \min_{k,j}(|g_k - g_j|). \tag{B.6}$$

Parameter $C_h = 1/15$ was chosen according to [111]. Note, that the step size could also be determined by a more elaborate scheme, updating the step over the runtime of the algorithm to potentially improve the complexity and accuracy of the algorithm. Further studying these details was deemed out of scope for this thesis, thus this straightforward definition is used without further justifications. Initial orientations $\varphi_k = 0$, $\varphi_k = \pi/2$, $\varphi_k = \pi$, $\varphi_k = 3\pi/2$ are set according to the corresponding g_k being on the lower, right, upper or left boundary respectively. The number of phase jumps is K, which ideally should be equal to the number of discrete eigenvalues in the discrete nonlinear Fourier domain of the pulse. The function isJump(x, y) returns "true" if $\arg(a(x)) \cdot \arg(a(y)) < 0$ and $|\arg(a(x)) - \arg(a(y))| > 1.3\pi$ are fulfilled. Otherwise it returns "false". Note that in addition to the steps shown in algorithm 3, it has to be assured that the while loop is exited after some maximum number of iterations or after the algorithm leaves some reasonable search area. Otherwise the algorithm might get stuck if e.g. the received pulse is too heavily distorted. A confined search area could be defined by an extended version of domain G given by $G = [L_\lambda - C_b \cdot h_\lambda, R_\lambda + C_b \cdot h_\lambda] \times [0 - C_b \cdot h_\lambda, U + C_b \cdot h_\lambda]$ where $C_b = 2$. While the presented algorithm can potentially be improved by a more refined definition of the search area, the above definition was used over the course of the thesis.

B.4. Fiber Parameter Estimation Algorithm

Fig. B.1 shows the flow-chart of the fiber parameter estimation algorithm discussed in section 4.10.2.

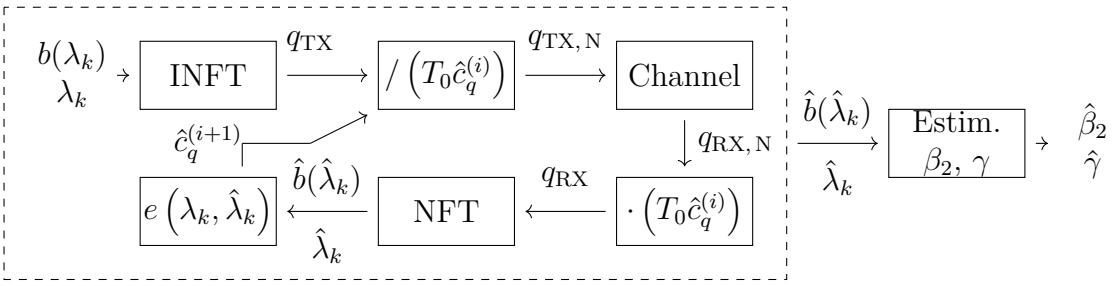


Figure B.1.: Flow chart of fiber parameter estimation with multi-solitons.

In Fig. B.1, the first two blocks correspond to the transmitter side, while the lower row of blocks on the left corresponds to the receiver side. The block outside of the dashed area is executed once a suitable estimate $\hat{c}_q$ was found by the iterative process

depicted inside the dashed area. Parameters $\hat{c}_q$, $\hat{b}(\hat{\lambda}_k)$, $\hat{\lambda}_k$, $\hat{\beta}_2$ and $\hat{\gamma}$ are estimates of c_q, $b(\lambda_k)$, λ_k, β_2 and γ respectively and $q_{\mathrm{TX/RX}}$ is the normalized transmitted/received signal while $q_{\mathrm{TX/RX,\,N}}$ are the denormalized quantities.

C

Additional Parameter Tables

In this section, additional parameter tables are given. These tables were removed from the main body of the thesis, but still are shown here for the sake of completeness. Note, that all the tables given hold parameters that more or less represent standard SSMF systems with coherent receivers. The deviations between tables mostly stem from the fact that the results in this thesis were obtained over several years and values were slightly altered from time to time to e.g. match the values found in some literature or to be able to more easily compare data with collaborators. None of the value changes, apart from things like e.g. the transmission distance or amplification type yield major differences in the system behavior.

C.1. Table for WDM Simulations

In the following, system parameters for the results presented in section 3.5 are given. Further parameters, which were changed in the conducted simulations, are given in the description of the respective simulation.

The number of propagating channels is given by N_{ch}, including the evaluated COI. Parameter N_{s} gives the number of samples per symbol, R_{Baud} is the systems symbol rate, β_{RO} is the roll-off factor of the RRC-pulses used, B_{G} is the guard-band between multiplexed channels and F_{n} is the noise figure of the used EDFA amplifiers. The fiber parameters are attenuation α, GVD parameter β_2 and nonlinearity parameter γ. The systems central wavelength is λ_{s}. The parameter ϕ_{NL}^{max} is used for the step-size

Parameter	Value	Parameter	Value
N_{ch}	5	α	0.2 $\frac{\mathrm{dB}}{\mathrm{km}}$
N_{s}	8	β_2	$-2.1173 \cdot 10^{-26}$ $\frac{\mathrm{s}^2}{\mathrm{m}}$
R_{Baud}	10 GBd	γ	1.3 $\frac{1}{\mathrm{W \cdot km}}$
β_{RO}	0.25	λ_{s}	1550 nm
B_{G}	2.5 GHz	ϕ_{NL}^{max}	10^-4
F_{n}	5 dB		

Table C.1.: Common parameters for WDM Simulations in section 3.5.

calculation using the nonlinear phase-method described in section 3.1.5.

C.2. Table for PLA-Enhanced Transmission Systems

In this table, the parameters for the simulations conducted on PLA-enhanced transmission systems are shown. Apart from fiber parameters, the parameters for DRA and EDFA amplification are given as well.

Parameter	Value	Parameter	Value
γ	≈ 1.3174 $\mathrm{W^{-1}km^{-1}}$	w_{spot}	5.0463 μm
D	17 ps/nm/km	T_{fib}	290 K
α	0.2 dB/km	α_{P}	0.25 dB/km
λ_{s}	1.55 μm	F_{n}	5
A_{eff}	$\approx 8.00011 \cdot 10^{-11}$ m^2		

Table C.2.: Simulation parameters for results from section 4.9 and appendix D.4.

C.3. Table for Fiber Parameter Estimation with Multiple Discrete Eigenvalues

In this table, the values for the simulations conducted on fiber parameter estimation, using pulses with multiple discrete eigenvalues, are given. This includes the parameters of the chosen probing pulses.

Note, that the third discrete eigenvalue $\lambda_3 = 0.15i$, was not used for the final estimation of β_2 and γ, but was utilized in the estimation of c_{q}. This is due to the fact that the algorithms implemented is relying on a non-zero real part of the discrete eigenvalue to compute c_{z}.

Parameter	Value
N_{span}	20
$\mathcal{L}_{\text{span}}$	80km
$\mathcal{L}$	1600km
β_2	$-2.117 \cdot 10^{-26}\text{s}^2\text{m}^{-1}$
γ	$1.2 \cdot 10^{-3}(Wm)^{-}1$
K	3

Parameter	Value
c_{q}	$2.38 \cdot 10^{11}$ $\text{W}^{-1/2}\text{s}^{-1}$
c_{z}	$1.058 \cdot 10^{-26}$ s^2m^{-1}
$[\lambda_1, q_{\text{d}}(\lambda_1)]$	$[-0.1 + 0.3i, 1.0 + 1.0i]$
$[\lambda_2, q_{\text{d}}(\lambda_2)]$	$[0.1 + 0.3i, 4.0 + 4.0i]$
$[\lambda_3, q_{\text{d}}(\lambda_3)]$	$[0.15i, 1.0i]$

Table C.3.: Simulation parameters for results from section 4.10.2 and appendix D.5.

C.4. Table for EM-Clustering Results

For the evaluation of the EM-clustering algorithm, the parameters slightly differ from the standard parameter set given in table C.1. The parameters which remain fixed in the conducted simulations are given in table C.4.

Parameter	Value
β_2	$-21.683\,\text{ps}^2/\text{km}$
γ	$1.3174\,\text{W}^{-1}\text{km}^{-1}$
λ_{s}	1550 nm
α	0.2 dB/km
α_{p}	0.25 dB/km
$P_{\text{P}}(0)$	18.095 dBm

Parameter	Value
T_{S}	8.0535
p	99.975%
R_{Baud}	1.5 Gbaud
B_{S}	58.9 GHz
$\mathcal{L}_{\text{span}}$	25 km

Table C.4.: Parameters used for clustering test in section 4.12.

Here, α is the fiber attenuation for the signal and α_{p} is the fiber attenuation for the pump lasers. The values are different, due to the fact that the signal and pump are situated at different wavelengths and thus experience differing losses. Parameter $P_{\text{P}}(0)$ is the initial pump power, T_{S} is the symbol duration in soliton units, p is the in-window energy factor and $\mathcal{L}_{\text{span}}$ is the span-length of the system used in the simulations in section 4.12.

D

Additional Simulation Results

In this section, additional results that did either did not fit into the main part of the thesis or are just variations of results already shown in the thesis, are presented. While the presented results will not be analyzed with the same rigor as the results in the main chapters, they are still deemed valuable enough to be discussed shortly. Where applicable, there will also be references to the respective sections the results belong to.

D.1. Additional Simulation Results for WDM Systems

In the following, additional simulation results for the MI of the WDM system presented in section 3.5 are given. They are variations of the simulation results shown in section 3.5.3. If not stated otherwise, the parameter set from table C.1 is used. Parameters that vary between simulations are stated in the captions of the respective simulations. Similar to Fig. 3.15, simulation results for other span-lengths were obtained. Figs. D.1 and D.2 show the MIs for the respective point-to-point systems, while keeping the overall transmission distance identical. The same set of parameters was tested for the network case and the results are shown in Figs. D.3 and D.4. Note, that the shown curves are either obtained using the CDC algorithm or the DBP algorithm with a varying number of steps per span N_{dbp} for equalization at the receiver. While the linear signal distortions can be mitigated by both algorithms, the nonlinear distortions can only be mitigated by using the DBP algorithm. A more detailed discussion of this

topic can be found in section 3.5.2.

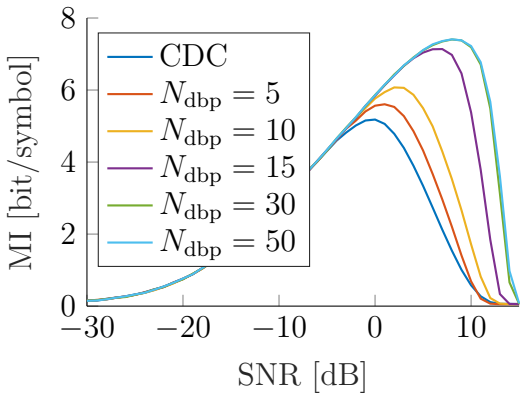


Figure D.1.: 16x16 MPSK transmission over 25x80km (Point-to-Point).

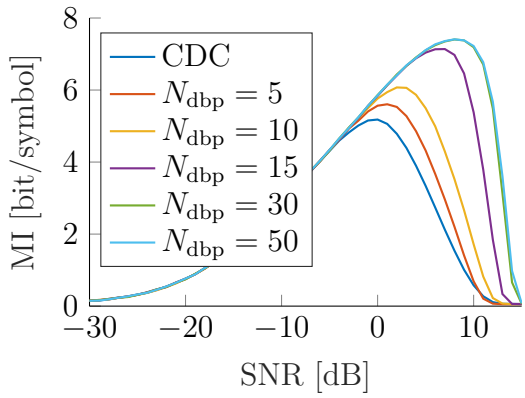


Figure D.2.: 16x16 MPSK transmission over 40x50km (Point-to-Point).

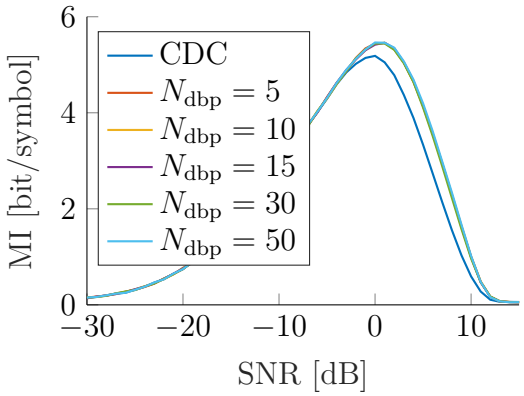


Figure D.3.: 16x16 MPSK transmission over 25x80km (Network Case).

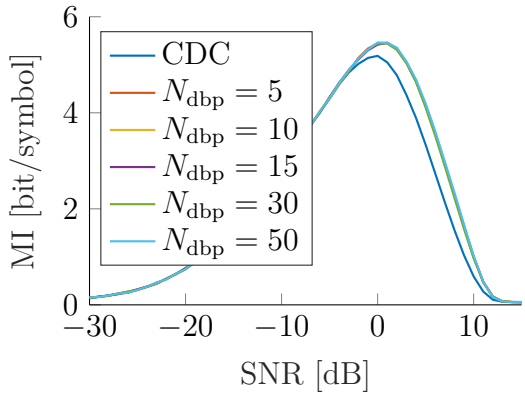


Figure D.4.: 16x16 MPSK transmission over 40x50km (Network Case).

D.2. Accuracy of the Continuous Spectrum NFT

The results shown in Fig. D.5, are obtained using the same simulation setup as described in section 4.4.1. The simulations differ only in the pulse used to test the accuracy of the algorithms. In Fig. D.5 the amplitude of the rectangular pulse was set to $A = 2$, while in Fig. 4.8 from section 4.4.1 in the main part of the thesis the amplitude was set to $A = 6$. Note, that the RK is not visible in Fig. D.5. This is

due to the fact that it is very close to the results obtained for TD and thus is almost fully covered by the other curve. There are also results for the accuracy of the NFTs

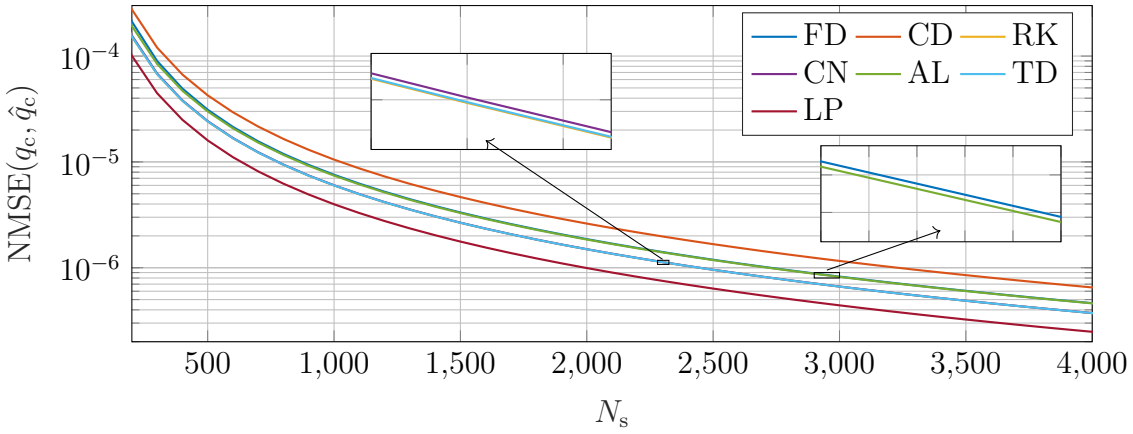
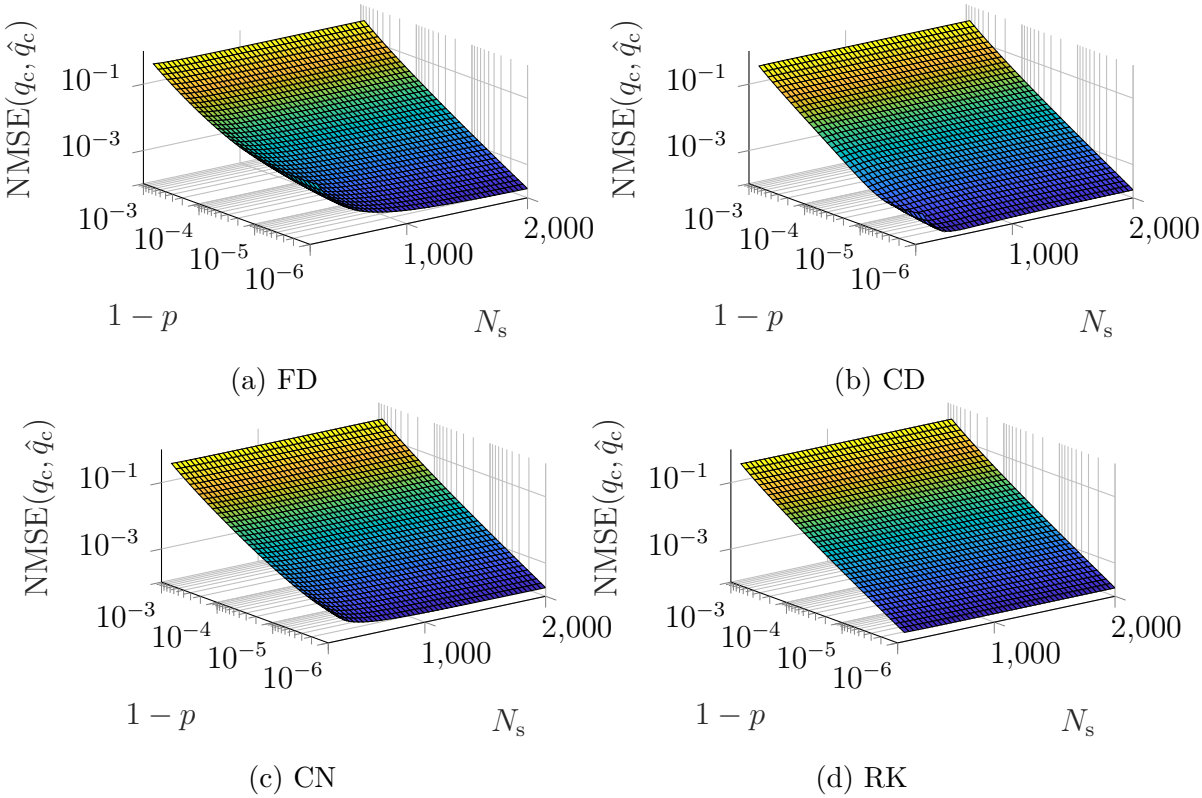


Figure D.5.: Accuracy of continuous spectrum estimation for the rectangular pulse ($A = 2$, $T = 1$).

which were computed using the SY-pulse as a test pulse. The results for several of the algorithms presented in 4.4.1 are given in Figs. D.6 and D.7. It can be seen, that at least when speaking about orders of magnitude, the algorithms evaluated in Figs. D.6 and D.7 all react similar to changes in the in-window energy factor p, except for cases where N_s is very low. The number of samples N_s decreses the NMSE for higher values for some algorithms, but as visible in e.g. Fig. D.6d and Fig. D.7c, some algorithms perform more or less independent of N_s for the SY-pulse. Without going into detail too much, the results for the SY-pulse can be compared with the results fo the rectangular pulse from Figs. 4.8 and D.5. The different behavior in regards to N_s could be due to the fact that the SY-pulse is a more smooth function than the rectangular pulse and thus increasing the resolution at the transients of the rectangular pulse always improves the estimation, while the resolution is no major concern for the transformation of the SY-pulse after some threshold value is reached. Increasing the value p improves the accuracy for all algorithms tested with the SY-pulse. This could be due to the fact that the SY-pulse is infinitely broad and thus is always truncated for finite temporal support. Contrary to that, the rectangular pulse can be fully supported on a finite time-window, which is why the in-window energy was not varied for the accuracy test of the rectangular pulse. If the time-window would be decreased beyond the time-support, the pulse-duration would be altered and the influence of this is well captured in the closed-form expressions for the nonlinear



(a) FD (b) CD

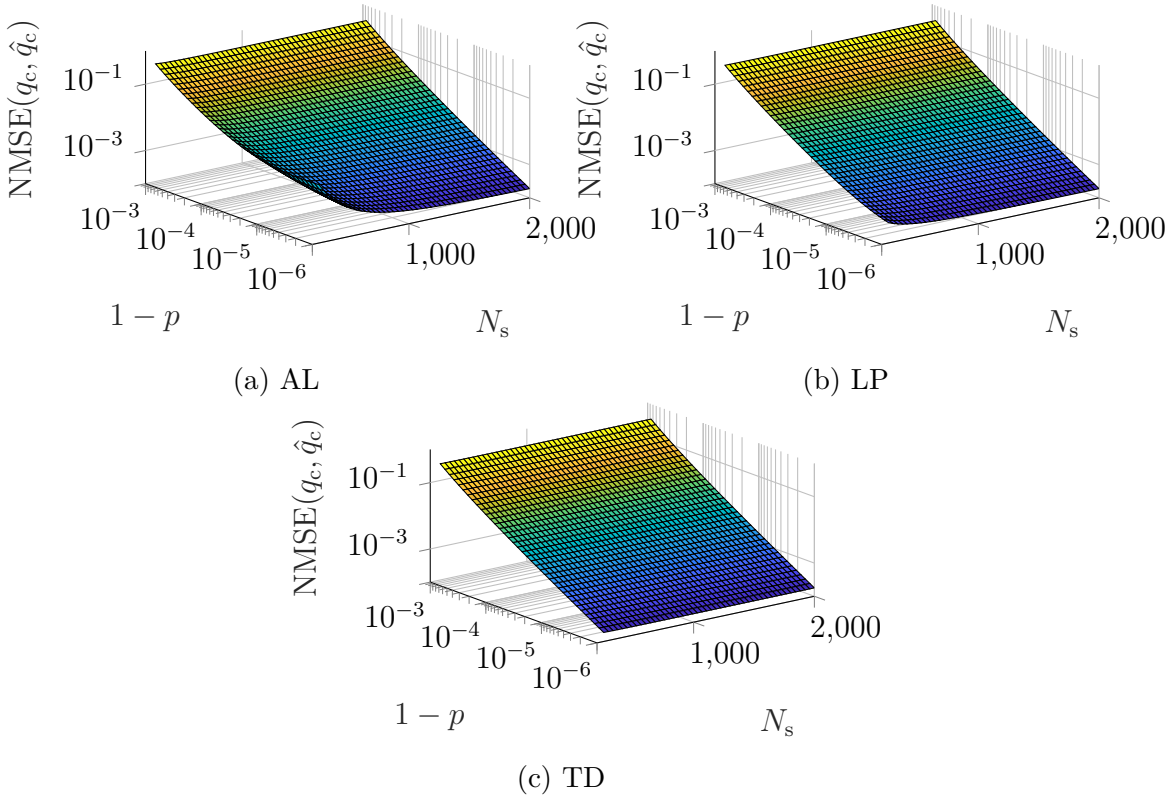
(c) CN (d) RK

Figure D.6.: Accuracy of the continuous spectrum estimation of a SY-pulse (Part 1).

Fourier transform of the rectangular pulse given in section 4.3.1.

D.3. Accuracy of the Eigenvalue Removal NFT

In section 4.4.3, the results for $K = 2$ discrete eigenvalues were already presented in Fig. 4.15. Similar results were also obtained for a transmission pulse generated with $K = 3$ discrete eigenvalues. The results are shown in Fig. D.8. The ranges for the two parameters changed during simulation are $A \in (0.6,\ 1.5)$ and $C_{\mathrm{EV}} \in (0.3,\ 1.2)$. Most of the observations made for $K = 2$ in section 4.4.3 are also true for $K = 3$. Some additional remarks on the $K = 3$ case are given in the following. It can be seen, that for the case shown in Fig. D.8 the continuous spectrum detection is not benefiting from the truncation process in most cases. The algorithm can potentially even degrade the quality of the obtained continuous spectrum, while in a small region there is also some improvement visible which it is quite confined. No further studies were conducted to understand under which conditions this improvement can be achieved in a more



(a) AL

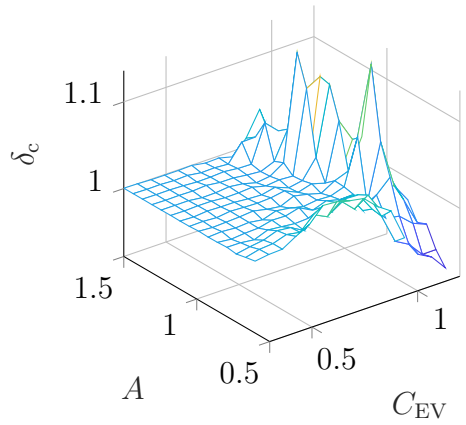
(b) LP

(c) TD

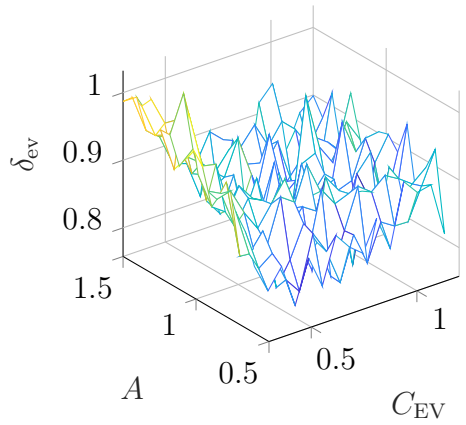
Figure D.7.: Accuracy of the continuous spectrum estimation of a SY-pulse (Part 2).

generalized way. Thus a similar approach to splitting the detection into two parallel routines such as already discussed for the case $K = 2$ seems beneficial in this case as well.

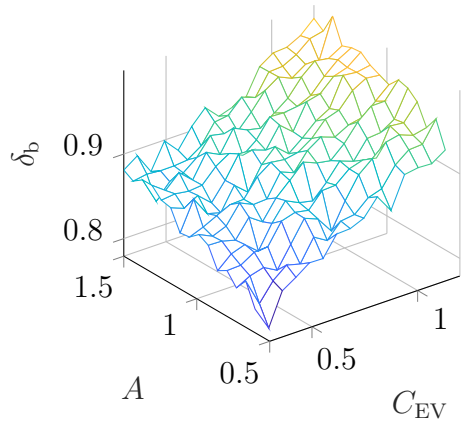
The detection accuracy of discrete eigenvalues and $b(\lambda_k)$ values can be seen in Figs. D.8b and D.8c. Compared to the $K = 2$ case, the improvement over purely search-based detection is less pronounced, but still the removal based algorithm is generally more accurate. For the discrete eigenvalues (Figure D.8b) the algorithm still shows some improvement, but there is no clear slope visible. One possible explanation could be that for $K = 3$ the discrete eigenvalues have a larger imaginary part and therefore, the truncation threshold T_c could be reached more quickly. As a result the numerical error could not be reduced further by truncation. Further, since T_c poses an ultimate threshold for the reduction in numerical complexity, the current implementations complexity will be very close to the complexity of the search-based method if T_c is very close to the temporal width of the received pulse. In Figure D.8c one can see that for the b-values the improvement is reduced for larger C_{EV} and A. We assume, that the



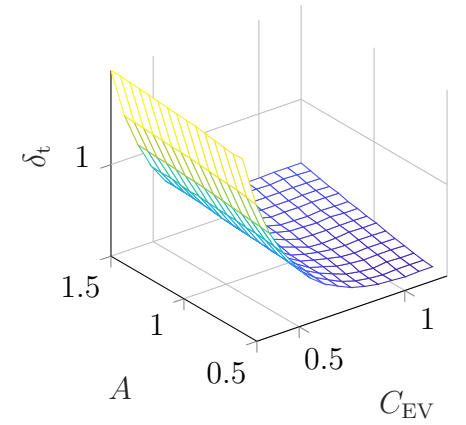
(a) Continuous spectrum NMSE quotient δ_c.



(b) Discrete eigenvalues NMSE quotient δ_{ev}.



(c) Discrete b-values NMSE quotient δ_b.



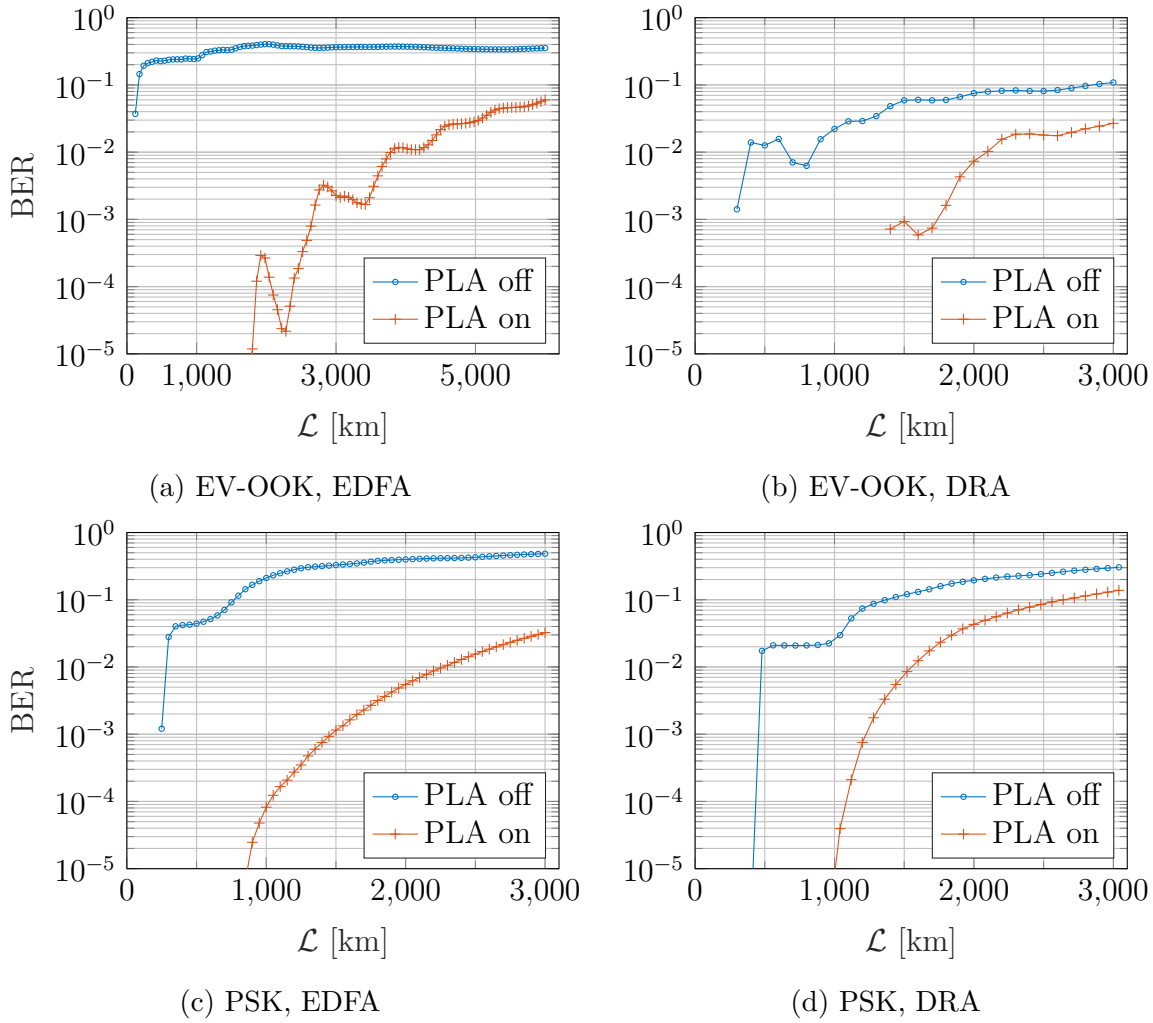
(d) Computation time quotient δ_t.

Figure D.8.: Comparison of detection algorithms for $K = 3$ eigenvalues.

reason for this behaviour is the less accurate detection of large discrete eigenvalues, which might in return cause larger errors in the eigenvalue removal step or even fail to remove all eigenvalues from the pulse. As discussed earlier in section 4.4.3, the algorithm defaults to the accuracy of the search-based NFT, if all eigenvalue removal steps fail. In Fig. D.8d, the difference in computation time is depicted and it can be seen, that for small values of C_{EV}, the ER-NFT takes longer to compute than the purely search-based method. This effect is much less pronounced for the case of $K = 2$.

D.4. PLA-Enhanced Multi-Soliton Systems

For the PLA-enhanced transmission described in section 4.9, further results were obtained for various modulation schemes. Since they give similar insights into the effectivity of the PLA-approach, they are only shortly presented and discussed in this section for the sake of completeness. Apart from the MPSK modulation case discussed in section 4.9, further results were obtained for EV-OOK and PSK modulation systems. The results for these are depicted in Fig. D.9. The span-length and Baud-rate are



(a) EV-OOK, EDFA

(b) EV-OOK, DRA

(c) PSK, EDFA

(d) PSK, DRA

Figure D.9.: BER curves for PLA-enhanced systems

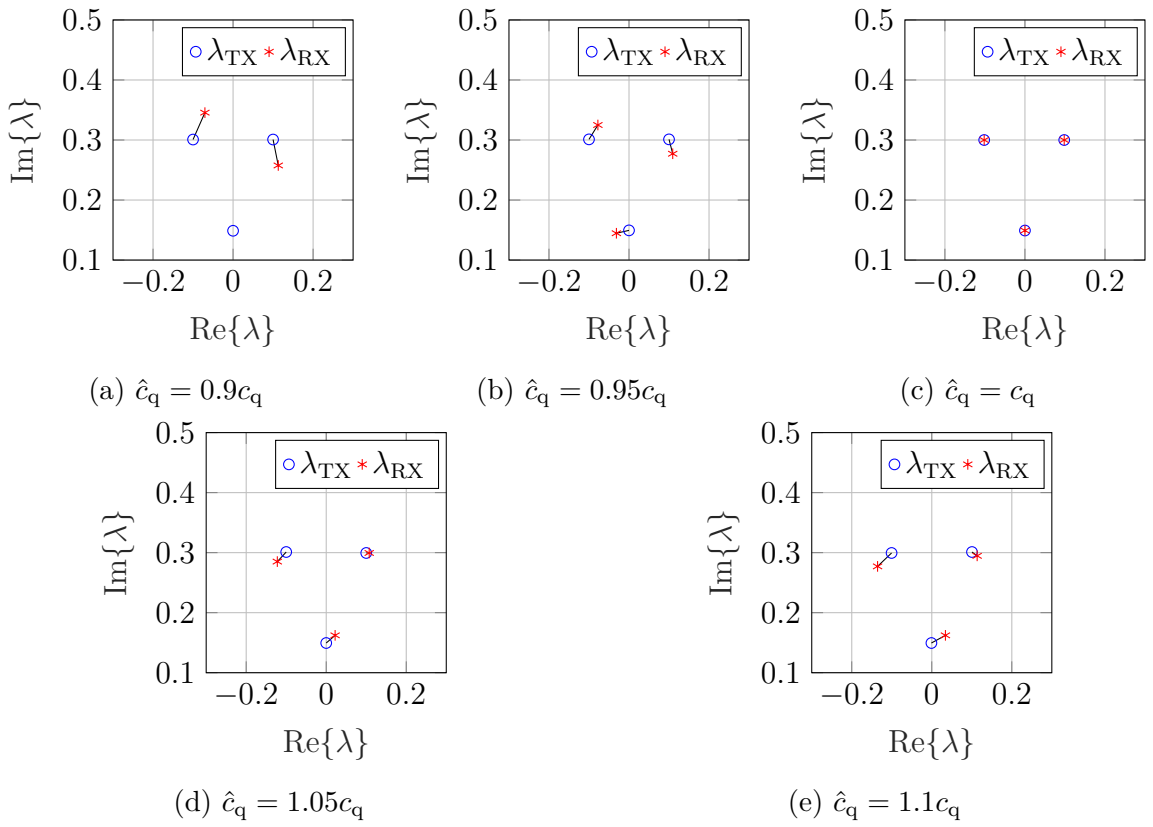
$\mathcal{L}_{\text{span}} = 60$ km, $R_{\text{Baud}} = 0.6$ GBaud for the EDFA amplified EV-OOK case presented in Fig. D.9a, $\mathcal{L}_{\text{span}} = 100$ km, $R_{\text{Baud}} = 0.57$ GBaud for the DRA amplified EV-OOK

case in Fig. D.9b, $\mathcal{L}_{\text{span}} = 10$ km, $R_{\text{Baud}} = 0.8$ GBaud for the EDFA amplified PSK case in Fig. D.9c and $\mathcal{L}_{\text{span}} = 80$ km, $R_{\text{Baud}} = 1.2$ GBaud for the DRA amplified PSK case in Fig. D.9d. The system parameters used are identical to the ones used in section 4.9 and can be found in table C.2. It can be seen, that for all cases shown in Fig. D.9, the PLA approach can improve the systems BER significantly. For EDFA amplification and EV-OOK signaling the results presented in Fig. D.9a show that for the used span-length is is not feasible to transmit data at the used Baud-rate without using the PLA approach. For the results using PLA rescaling it can be seen that, just as a quick check using $10^{-}3$ as a FEC-threshold, transmission distances beyond 2500 km can be achieved. The unsteady movement of the "PLA on" curve could be due to the choice of the temporal support, causing the symbol to expand beyond the allocated window for certain transmission distances. This, however, was not further studied. For the EDFA case, using PSK modulation in Fig. D.9c, the improvement in the PLA-enhanced case is rather significant as well. Note, that in this case the effect from Fig. D.9a is not visible, which might be due to a better choice for the temporal support. For both DRA cases, presented in Figs. D.9b and D.9d, the improvement is less impressive, but still quite significant. Again there are some fluctuations visible in D.9b in both curves. This further strengthens the conjecture that some of the used pulses extended beyond their allotted time-window for certain transmission distances. Using the PSK modulation scheme yields a more uniform pulse-set than the EV-OOK method and thus it is more likely that the heuristically chosen temporal support for the simulations depicted in Figs. D.9a and D.9c caused the visible fluctuations in the BER.

D.5. Fiber Parameter Estimation

In addition to the results presented in section 4.10, additional simulation results were obtained for the ideal channel case and the PLA-enhanced system case. The parameters used are identical to the parameters used in section 4.10.2 and are given in table C.3. To visualize the signal degradation due to normalization parameter $\hat{c}_{\text{q}}$ diverging from the ideal value c_{q}, two channel configurations were tested. First the channel was assumed to behave according to the ideal channel model, enabling the solitonic pulses to propagate ideally through the fiber if normalized correctly. This normalization and denormalization process then was slightly disturbed by setting $\hat{c}_{\text{q}} = C \cdot c_{\text{q}}$ with some

$C \neq 1$. The movement of the discrete eigenvalues in the positive complex halfplane due to this disturbance is shown in Figs. D.10a - D.10e. Note, that for Fig. D.10a

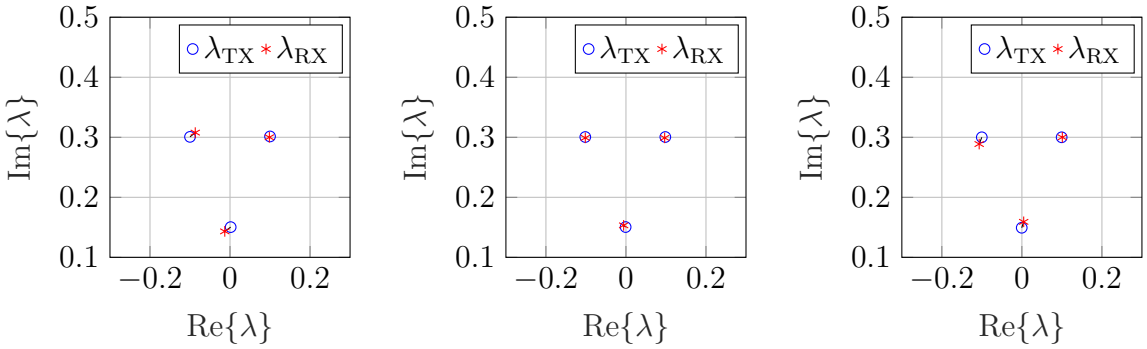


(a) $\hat{c}_\mathrm{q} = 0.9c_\mathrm{q}$ (b) $\hat{c}_\mathrm{q} = 0.95c_\mathrm{q}$ (c) $\hat{c}_\mathrm{q} = c_\mathrm{q}$

(d) $\hat{c}_\mathrm{q} = 1.05c_\mathrm{q}$ (e) $\hat{c}_\mathrm{q} = 1.1c_\mathrm{q}$

Figure D.10.: Transmitted and received discrete eigenvalues in parameter estimation setup for the ideal channel.

the discrete eigenvalue with the smallest imaginary part could not be detected at the receiver. Further it can be seen that the discrete eigenvalues are perturbed the least for the case $\hat{c}_\mathrm{q} = c_\mathrm{q}$. This is not necessarily true for channels with e.g. an EDFA amplification scheme, as depicted in Fig. 4.32. In Figs. D.11a - D.11c, the same parameter estimation algorithm is tested with an EDFA amplified fiber channel. Here it can be seen that, once more, the deviations are more pronounced if $C \neq 1$. Additionally, the ideal case depicted in D.11b now also shows some deviations. This is due to the fact that the PLA scheme does not fully counteract the perturbations due to attenuation. This effect is explored in more detail in section 4.10.1, albeit only for the fundamental soliton case.



(a) $\hat{c}_\mathrm{q} = 0.975 c_\mathrm{q}$ (b) $\hat{c}_\mathrm{q} = c_\mathrm{q}$ (c) $\hat{c}_\mathrm{q} = 1.025 c_\mathrm{q}$

Figure D.11.: Transmitted and received discrete eigenvalues in parameter estimation setup for PLA-enhanced EDFA transmission.

D.6. EV-OOK Labeling Schemes

Apart from the results shown in Figs. 4.36a and 4.36b, further results on the impact of the "gray" labeling algorithm presented in section 4.11 were obtained. Some of the results will be presented in the following. Note, that the systems simulated in section 4.11 are not identical to the system models for this section. Thus, while the results still show similar trends, they can not be compared directly. The used system parameters are given in table C.2. In Fig. D.12, the BER for transmission over a DRA amplified fiber channel is given for "gray" and binary labeling. Additionally, the

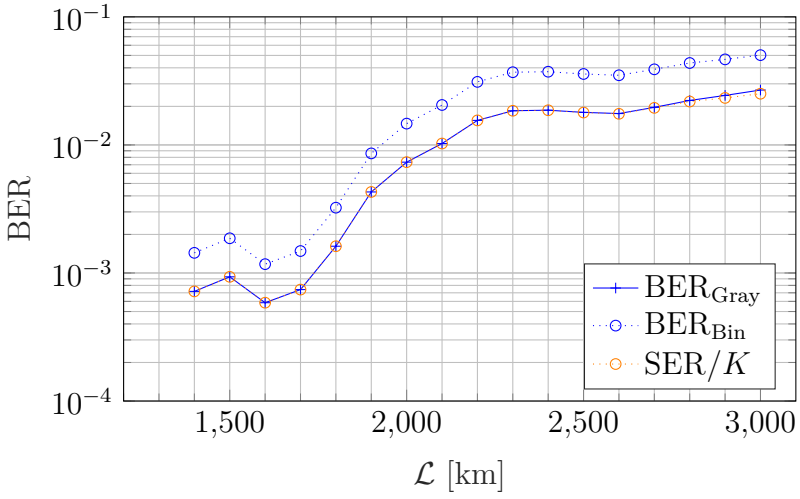


Figure D.12.: BER curves for DRA and EV-OOK with PLA approach. ($R_\mathrm{Baud} = 0.57$ GBaud, $\mathcal{L}_\mathrm{span} = 100$ km, $K = 5$)

SER divided by the number of discrete eigenvalues $K = 5$ is given. This is the lower bound for which only one bit-error per symbol-error occurs. As it can be seen, similar to the results from section 4.11, this bound is achieved for 'gray' labeling for a fairly long distance. Around 2800-2900km the results start to diverge from the optimum value, which means that the heuristic channel model assumed in section 4.11 is no longer adequately modeling the behavior of the simulated channel. Similarly, results for an EDFA amplified fiber channel are shown in Fig. D.13. It can be seen, that

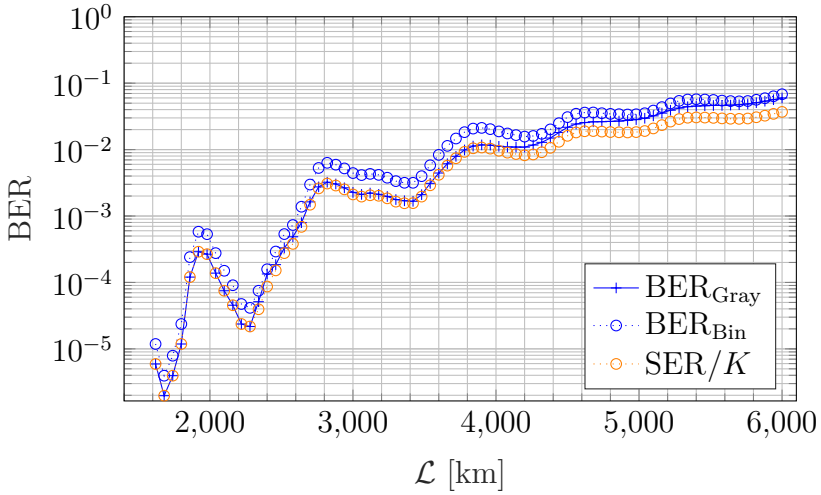


Figure D.13.: BER curves for EDFA and EV-OOK with PLA approach. ($R_{\mathrm{Baud}} = 0.6$ GBaud, $\mathcal{L}_{\mathrm{span}} = 60$ km, $K = 5$)

the behavior of the 'gray' labeled curve is similar to the one in Fig. D.12, but the deviation gets even more pronounced for longer distances. For $\approx$ 5400-6000km the values of $\mathrm{BER}_{\mathrm{Gray}}$ is similar to $\mathrm{BER}_{\mathrm{Bin}}$, which means that the heuristic model does no longer hold in this case as well.

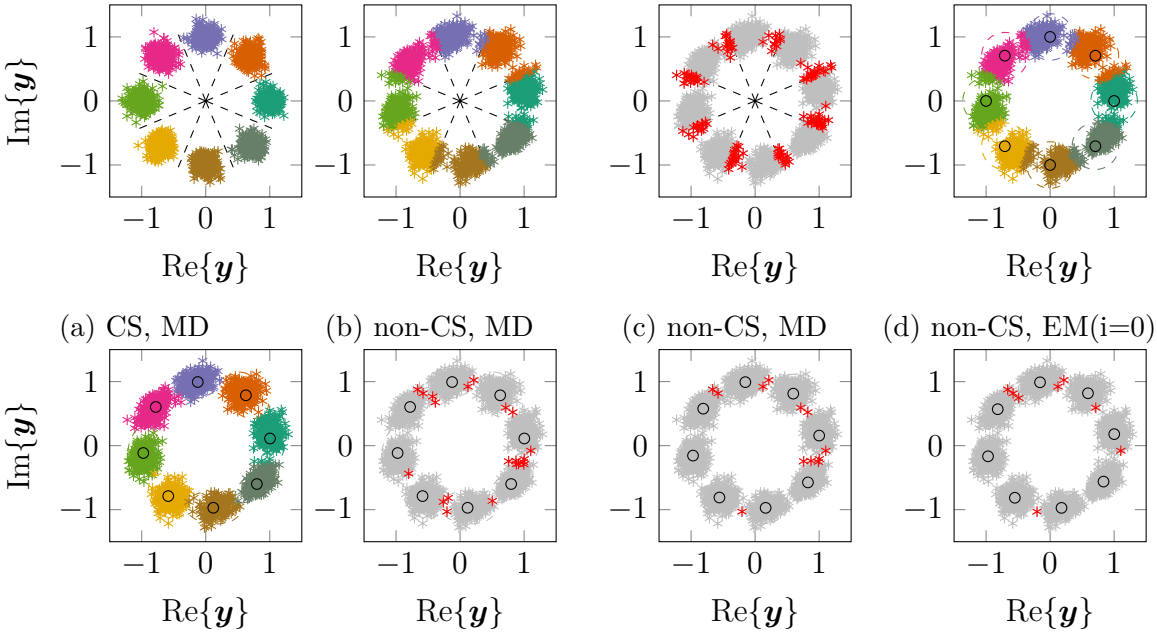
D.7. Clustering Algorithm for Discrete Spectrum Data

In the following, some additional results and figures regarding the topic of clustering, discussed in section 4.12, will be presented. The two main groups of additional material contain an example for the iterative grouping of a data set using the EM algorithm and some additional results for the performance of transmission systems using the k-means (KM) algorithm for detection at the receiver. Note, that the KM algorithm

is not discussed in detail in this thesis. A detailed treatment of the KM algorithm can be found in e.g. [147].

D.7.1. Example for Clustering with the EM Algorithm

To illustrate the iterative process by which the EM-clustering algorithm is grouping the received data $\boldsymbol{y}$, a simplified example of a 8-PSK modulated signal is used. The channel will just add noise before the receiver. The noise itself is AWGN which either is circularly symmetric (CS) of non-CS. Several depictions of the received constellations are given in Fig. D.14. In Fig. D.14a, the received constellation is superimposed by CS



(a) CS, MD (b) non-CS, MD (c) non-CS, MD (d) non-CS, EM(i=0)

(e) non-CS, EM(i=1) (f) non-CS, EM(i=1) (g) non-CS, EM(i=2) (h) non-CS, EM(i=6)

Figure D.14.: Received constellations for MD and EM detection.

AWGN. The black dotted lines show the decision boundaries for MD detection. It can be seen, that as long as the noise is CS, the MD boundaries are ideal. On the contrary, in Fig D.14b, the superimposed noise is non-CS and it can be seen, that the rotation and deformation of the symbol clouds leads to suboptimal detection by MD. In Fig. D.14c, the same data is shown, but the errors are marked in red. The resulting SER for the data and detection scheme shown in Figs. D.14a - D.14c is SER $= 7.00 \cdot 10^{-2}$. If now EM-clustering is used, the algorithm tries to find parameters to best represent

the received data by a MOG. In Fig. D.14d, these Gaussians for iteration $i = 0$ are shown. The black circles give their initial mean values, set by the transmission alphabet. The dashed circles visualize the 95% confidence interval for each Gaussian. In the next iteration $i = 1$, new mean values are computed for the clusters. The mean values are again given by the black circles in Fig. D.14e. Also the data is clustered again using the new values for mean and (co-)variance of the Gaussians. It can be seen, that already after the first iteration less detection errors occur. In Fig. D.14e this is especially visible for the yellow cluster in the bottom left. The errors are also visualized in Fig. D.14f by the red data points. The error after the first iteration is reduced to SER $= 1.05 \cdot 10^{-2}$. Similarly the errors after iteration $i = 2$ and $i = 6$ are depicted in Figs. D.14g and D.14h and result in error rates of SER $= 6 \cdot 10^{-3}$ and SER $= 4 \cdot 10^{-3}$ respectively. It can be seen, that in this case the EM algorithm reduced the SER by more than one order of magnitude in a few iterations.

D.7.2. Results for K-Means Clustering

KM also is an algorithm for clustering data. Compared to EM clustering the algorithm is easier to implement, but in many cases also is less good at clustering data. A more detailed treatment of KM clustering can be found in e.g. [147]. In Fig. D.15 the BER for transmission of a PSK modulated signal over a DRA amplified fiber channel with varying transmission distances $\mathcal{L}$ is shown. The modulated quantities are either

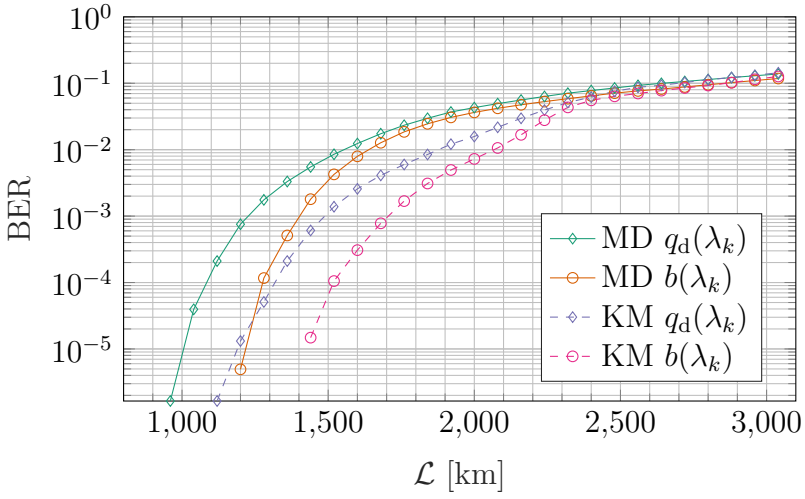


Figure D.15.: BERs for DRA amplification and 8-PSK. ($R_\mathrm{Baud} = 1.2$ GBaud, $\mathcal{L}_\mathrm{span} = 80$ km $K = 2$)

the discrete nonlinear spectral amplitudes $q_\mathrm{d}(\lambda_k)$ or the $b(\lambda_k)$ values. The detection is either done by MD or KM clustering. It can be seen, that in both cases clustering by KM significantly improves the BER in a certain range. Assuming the FEC-threshold 10^{-3}, in both cases the achievable transmission distance can be increased by ≈ 300 km.

Similar results were obtained for the EDFA amplified channel. The results are given in D.16. It can be seen, that the influence of using KM clustering on the BER is much

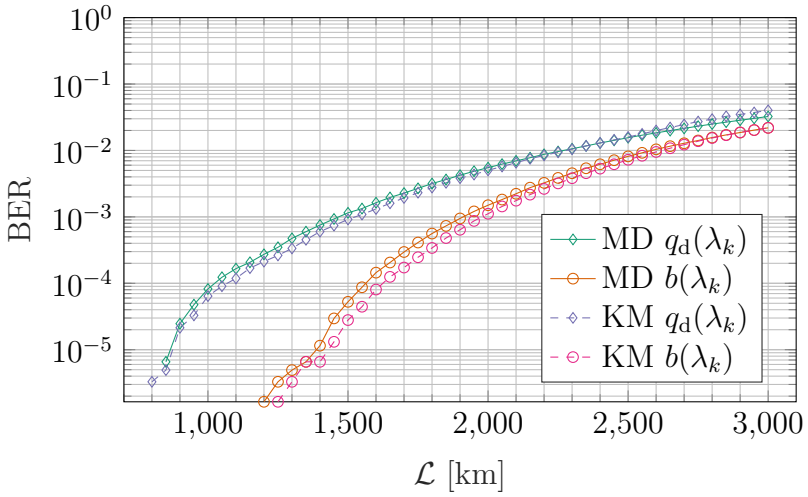


Figure D.16.: BERs for EDFA amplification and 8-PSK. ($R_\mathrm{Baud} = 0.8$ GBaud, $\mathcal{L}_\mathrm{span} = 10$ km $K = 2$)

smaller in this case. While this was not studied further, this decrease in effectivity could be caused by the reduced influence of nonlinear distortions. Compared to the DRA case, the fluctuation in signal power over each span of EDFA amplified fiber is much more pronounced. Thus, for large portions of the transmission distance, the signal power is rather low and signal distortions are more likely to stem from fiber attenuation and the resulting mismatch between the theoretic and the simulated channel model.

E

Notation, Common Functions, Symbols and Acronyms

E.1. Notation and Functions

$(\cdot)^*$ Complex conjugate

$(\cdot)^{\mathrm{H}}$ Hermetian of vector/matrix

$(\cdot)^{\mathrm{T}}$ Transposition of vector/matrix

$\Im\{\cdot\}$ Imaginary part of complex number

$\Re\{\cdot\}$ Real part of complex number

$\angle(\cdot)$ Argument of complex number

$\boldsymbol{X}$ Matrix containing entries x_{ij} with i, j denoting row and column respectively

$\boldsymbol{x}$ Vector containing entries $x_1, x_2, \cdots, x_N$

$\langle \cdot, \cdot \rangle$ Inner product

$\mathbb{E}[\cdot]$ Expected value

$\mathbb{H}(\cdot)$ Entropy of a random variable

$\mathbb{H}(\cdot|\cdot)$ Conditional entropy of two random variables

$\mathbb{I}(\cdot;\cdot)$ Mutual information

$\mathcal{F}\{\cdot\}$ Fourier transform

$\mathcal{F}^{-1}\{\cdot\}$ Inverse Fourier transform

$\mathcal{H}\{\cdot\}$ Hilbert transform

$\mathcal{H}^{-1}\{\cdot\}$ Inverse Hilbert transform

$\mathrm{Cov}[\cdot,\cdot]$ Covariance

$\mathrm{INFT}(\{\cdot,\cdot,\cdot\})$ Inverse nonlinear Fourier transform

$\mathrm{NFT}(\cdot)$ Forward nonlinear Fourier transform

$\mathrm{Var}[\cdot]$ Variance

$\mathrm{W}(\cdot,\cdot)$ Wronskian

$\mathrm{supp}(\cdot)$ Support of a random variable

$\sigma(\cdot)$ Spectrum of an operator

$|\cdot|$ Absolute value

$||\cdot||$ Norm

E.2. Symbols

A_{eff} Effective area

B_{G} Guard bandwidth

B_{S} Symbol bandwidth

B_{rec} Receiver bandwidth

D_{PMD} PMD parameter

D Dispersion parameter

E_{χ} Complex electrical amplitude in eigenpolarization χ

E_{ψ} Complex electrical amplitude in eigenpolarization ψ

E_{x} Complex electrical amplitude in x-polarization

E_{y} Complex electrical amplitude in y-polarization

F_{eff} Effective noise figure for Raman amplifiers

F_{n} Noise figure

G_{EDFA} EDFA gain

$G_{\text{on-off}}$ On-off gain for Raman amplifiers

K_{T} Shifted phonon occupancy factor ($K_{\text{T}} = 1 + \eta_{\text{P}}$)

L_0 distance normalization coefficient (NFT)

$N_{\text{b,error}}$ Number of bit-errors

N_{b} Number of total transmitted bits

N_{ch} Number of channels

N_{dbp} Number of steps per span for DBP algorithm

N_{span} Number of spans in transmission link

N_s Number of samples

P_0 power normalization coefficient (NFT)

P_N Noise power

P_P Pump power

P_S Signal power

$Q(\tau, \ell)$ Complex envelope of a narrowband signal in natural units

$Q_m(\tau, \ell)$ Entries of signal vector $\boldsymbol{Q}(\tau, \ell)$

R_Baud Baud rate

T_0 temporal normalization coefficient (NFT)

T_S Symbol duration

T_fib Fiber temperature

$\bar{\tau}_{\mathrm{g},\chi}$ Group-delay in eigenpolarization χ

$\bar{\tau}_{\mathrm{g},\psi}$ Group-delay in eigenpolarization ψ

$\Delta\ell$ step-size (SSFM)

$\Gamma(\cdot)$ Complex gamma function

Ω_R Raman shift

α_NR Step modifier for Newton-Raphson search

α Fiber loss coefficient

$\bar{\beta}_2$ Average GVD for the SCME channel model

$\beta(\omega)$ Mode-propagation constant

β_2 Group velocity dispersion coefficient

β_χ Propagation constant of eigenpolarization χ

β_RO Roll-off factor

β_ψ Propagation constant of eigenpolarization ψ

β_x Propagation constant of x-polarization

β_y Propagation constant of y-polarization

β Propagation constant

$\boldsymbol{E}(\ell, \tau)$ Complex electrical amplitude vector

$\boldsymbol{Q}(\tau, \ell)$ Vector of complex envelopes of a multimode signal in natural units

$\boldsymbol{q}_{\mathbf{c}}(\lambda)$ Continuous nonlinear spectrum vector (multi-mode case)

$\boldsymbol{q}_{\mathbf{d}}(\lambda_k)$ Discrete nonlinear spectrum vector (multi-mode case)

$\boldsymbol{q}(t, z)$ Vector of complex narrowband signal envelopes in soliton units

$\boldsymbol{v}(\lambda, t, z)$ eigenfunctions of isospectral Lax operator

$\chi^{(3)}$ Cubic material polarization

ℓ Propagation distance in natural units

η_{P} Phonon occupancy factor

η imaginary part of a generalized frequency λ

γ Nonlinearity coefficient

$\hat{D}_{\mathrm{dbp}}$ Linear terms for DBP

$\hat{D}$ Linear terms of the NLSE

$\hat{N}_{\mathrm{dbp}}$ Nonlinear terms for DBP

$\hat{N}$ Nonlinear terms of the NLSE

$\hat{Q}(\omega, \ell)$ Fourier transform of a narrowband signal in natural units

κ Nonlinear coupling coefficient for the SCME channel model

λ_k Discrete eigenvalues

λ generalized frequency

λ Wavelength in natural units (only used in Section 3)

$\mathbf{L}(z)$ isospectral operator of Lax pair

$\mathbf{M}(z)$ spatial propagation operator of Lax pair

$\mathbf{P}(z)$ temporal propagation operator of Lax pair

$\mathbf{b}_{\mathrm{RX}}$ Received bitstream

$\mathbf{b}_{\mathrm{TX}}$ Transmitted bitstream

$\mathcal{H}$ Hilbert space

$\mathcal{L}_{\mathrm{eff}}$ Effective length

$\mathcal{L}_{\mathrm{span}}$ Length of one fiber-span

$\mathcal{L}$ Transmission distance

$\mathcal{Z}$ Normalized transmission distance

$\mathrm{diag}(\cdot)$ diagonal matrix operator

$\mathrm{d}\bar{\tau}_{\mathrm{g}}$ Differential group-delay

e Euler's number

j Imaginary unit $\mathrm{j} = \sqrt{-1}$

μ Mean of a random variable

ω_{c} Angular carrier frequency

ω Angular frequency

ϕ_{NL}^{max} Maximum allowed nonlinear phase-rotation in nonlinear phase-rotation method

ρ_{ME} Modulation efficiency

$\sigma^2_{\mathrm{ASE,span}}$ ASE noise variance per span

σ^2_{ASE} Total ASE noise variance

σ^2_{DRA} Noise variance for the DRA channel

σ^2_{IDRA} Noise variance for the IDRA channel

σ^2_{N} Noise variance

σ^2 Variance of a random variable

τ Temporal parameter in natural units

ξ real part of a generalized frequency λ

$\{a(\lambda), b(\lambda), \bar{a}(\lambda), \bar{b}(\lambda)\}$ Scattering coefficients

$\{\boldsymbol{\phi}(\lambda,t), \bar{\boldsymbol{\phi}}(\lambda,t), \boldsymbol{\psi}(\lambda,t), \bar{\boldsymbol{\psi}}(\lambda,t)\}$ Jost solutions for the NLSE-based NFT

$a(\lambda)$ Nonlinear Fourier coefficient

$b(\lambda)$ Nonlinear Fourier coefficient

$b_m(\lambda)$ Nonlinear Fourier coefficients for the multi-mode case

c Speed of light in vacuum

f_{c} Carrier frequency

f Frequency in linear Fourier domain

$g_{\mathrm{RRC}}(\tau)$ Root raised cosine

g_{R} Raman gain

$h(t)$ Linear channel impulse response

h_{P} Planck constant

k_{B} Boltzmann constant

$n(\omega)$ Frequency dependent refractive index

n_{g} Group refractive index

n_{sp} Spontaneous emission factor

$q(t,z)$ Complex envelope of a narrowband signal in soliton units

$q_{\mathrm{c}}(\lambda)$ Continuous nonlinear spectrum

$q_{\mathrm{c}}^{(m)}(\lambda)$ M-th continuous nonlinear spectrum (multi-mode case)

$q_{\mathrm{d}}(\lambda_k)$ Discrete nonlinear spectrum

$q_{\mathrm{d}}^{(m)}(\lambda_k)$ M-th discrete nonlinear spectrum (multi-mode case)

$q_{\mathrm{fs}}(t,z)$ Fundamental soliton in time domain

$q_{\mathrm{rec}}(t)$ Rectangular pulse in time domain

$q_{\mathrm{sy}}(t)$ Satsuma-Yajima pulse in time domain

$q_m(t,z)$ Entries of $\mathbf{q}(t,z)$

t Temporal parameter in soliton units

v_{g} Group-velocity

w_{spot} Spot radius

z Propagation distance in soliton units

$\mathbb{C}^-$ Negative complex half-plane

$\mathbb{C}^+$ Positive complex half-plane

$\mathbb{C}$ Set of complex numbers

$\mathbb{R}$ Set of real numbers

$\mathbb{Z}_0^-$ Set of negative integer numbers (including zero)

$\mathbb{Z}^-$ Set of negative integer numbers (excluding zero)

$\mathbb{Z}_0^+$ Set of positive integer numbers (including zero)

$\mathbb{Z}^+$ Set of positive integer numbers (excluding zero)

$\mathbb{Z}$ Set of integer numbers

E.3. Acronyms

AIR achievable information rate

AL Ablowitz-Ladik

AL-DLP Ablowitz-Ladik discrete layer peeling

ASE amplified spontaneous emission

AWGN additive white Gaussian noise

B2B back-to-back

BER bit error-rate

CD chromatic dispersion

CDC chromatic dispersion compensation

CDEP central difference eigenproblem

CDM central discretization method

CN Crank-Nicolson

COI channel of interest

CS circularly symmetric

DBP digital back-propagation

DCF dispersion compensating fiber

DFT discrete Fourier transform

DGD differential group-delay

DLP discrete layer peeling

DRA distributed Raman amplification

DSP digital signal processing

DT Darboux transform

EDFA erbium doped fiber amplifier

EM expectation maximization

ER eigenvalue removal

EV-OOK eigenvalue on-off keying

FB forward-backward

FD forward discretization

FEC forward error correction

FFT fast Fourier transform

FT Fourier transform

FWM four wave mixing

GLM Gelfand-Levitan-Marchenko

GVD group-velocity dispersion

I in-phase

IDRA idealized distributed Raman amplification

INFT inverse nonlinear Fourier transform

ISI inter symbol interference

IST inverse scattering transform

KM k-means

LP layer-peeling

MD minimum Euclidean distance

MF matched-filter

MI mutual information

MMF multi-mode fiber

MMI mismatched mutual information

MOG mixture of Gaussians

MPSK multi-ring phase-shift keying

MSE mean squared error

NFD nonlinear Fourier domain

NFDM nonlinear frequency division multiplexing

NFT nonlinear Fourier transform

NLSE nonlinear Schrödinger equation

NMSE normalized mean squared error

PAM pulse amplitude modulation

PDF probability density function

PJT phase jump tracking

PLA path-loss averaged

PMD polarization mode dispersion

PMF probability mass function

PPM pulse-position modulation

PSK phase-shift keying

Q quadrature

QAM quadrature amplitude modulation

RH Riemann-Hilbert

RK Runge-Kutta

ROADM reconfigurable optical add-drop multiplexer

RRC root raised cosine

RV random variable

RX receiver

SB Shannon-bound

SCME strong-coupling Manakov equation

SDM space division multiplexing

SE spectral efficiency

SER symbol error rate

SMF single mode fiber

SNR signal to noise ratio

SPM self-phase modulation

SRS stimulated Raman-scattering

SSFM split-step Fourier method

SSMF standard single mode fiber

SY Satsuma-Yajima

TBP time-bandwidth product

TBR time-bandwidth requirement

TD trapezoidal discretization

TX transmitter

VFS vectorial fundamental soliton

WDM wave division multiplexing

XPM cross-phase modulation

ZS Zakharov-Shabat

F

List of Publications

In the following, a list of all publications, that were written by the author during the work on this thesis, is presented. This does not include papers, that were written by the author before or after his doctoral studies. Posters at workshops/conferences and talks without a corresponding paper are not included. This list is just meant to give a quick overview of the authors publications. If they are referenced in the thesis, they are referenced by the identifier given in the bibliography section.

- Soliton Transmission with 5 Eigenvalues Over 2000km of Raman-Amplified Fiber
  - Authors: Leible, B.; Chen, Y.; Yousefi, M. I.; Hanik, N.
  - Conference/Journal: 20th International Conference on Transparent Optical Networks (ICTON), 2018
- Clustering Algorithm for Detection in the Discrete Nonlinear Fourier Spectrum
  - Authors: Leible, B.; García-Gómez, F. J.; Hanik, N.
  - Conference/Journal: 45th European Conference on Optical Communication (ECOC), 2019

- Eigenvalue Trajectories in Multispan Soliton Transmission Systems under Lumped and Distributed Amplification
  - Authors: Leible, B.; Hanik N.
  - Conference/Journal: 21st International Conference on Transparent Optical Networks (ICTON), 2019
- Approaches to Bit-Labeling for Eigenvalue On-Off-Keying Systems
  - Authors: Leible, B.; Göttsberger, T.; Hanik, N.
  - Conference/Journal: International Conference on Transparent Optical Networks (ICTON), 2020
- Stability of the Full Spectrum Nonlinear Fourier Transform
  - Authors: Leible, B.; Plabst, D.; Hanik, N.
  - Conference/Journal: International Conference on Transparent Optical Networks (ICTON), 2020
- Back-to-Back Performance of the Full Spectrum Nonlinear Fourier Transform and Its Inverse
  - Authors: Leible, B.; Plabst, D.; Hanik, N.
  - Conference/Journal: Entropy Journal 22 (10), 2020
- The Nonlinear Fourier Transform and its Extension to the Strong Coupling Multi-Mode Case
  - Authors: Leible, B.; Hanik, N.
  - Conference/Journal: 25st International Conference on Transparent Optical Networks (ICTON), 2023
- Algorithms for the Nonlinear Fourier Transform in the Strong Coupling Multi-Mode Case
  - Authors: Leible, B.; Hanik, N.
  - Conference/Journal: 25st International Conference on Transparent Optical Networks (ICTON), 2023

Bibliography

[1] G. J. Holzmann, "Data communications: The first 2500 years.," in *IFIP Congress (2)*, 1994, pp. 271–278.

[2] G. J. Holzmann and B. Pehrson, "The first data networks," *Scientific American*, vol. 270, no. 1, pp. 124–129, 1994.

[3] J. Hecht, *City of light: the story of fiber optics*. Oxford University Press on Demand, 2004.

[4] P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, "Fiber-optic transmission and networking: The previous 20 and the next 20 years," *Optics express*, vol. 26, no. 18, pp. 24 190–24 239, 2018.

[5] J.-M. Beaufils, "How do submarine networks web the world?" *Optical fiber technology*, vol. 6, no. 1, pp. 15–32, 2000.

[6] P. J. Winzer, "Optical networking beyond wdm," *IEEE Photonics Journal*, vol. 4, no. 2, pp. 647–651, 2012.

[7] M. Secondini, E. Agrell, E. Forestieri, D. Marsella, and M. R. Camara, "Nonlinearity mitigation in wdm systems: Models, strategies, and achievable rates," *Journal of Lightwave Technology*, vol. 37, no. 10, pp. 2270–2283, 2019.

[8] F. J. García-Gómez and G. Kramer, "Mismatched models to lower bound the capacity of dual-polarization optical fiber channels," *Journal of Lightwave Technology*, vol. 39, no. 11, pp. 3390–3399, 2021.

[9] F. J. García-Gómez and G. Kramer, "Regular perturbation and achievable rates of space-division multiplexed optical channels," in *2021 17th International Symposium on Wireless Communication Systems (ISWCS)*, IEEE, 2021, pp. 1–5.

[10] R. W. Tkach, "Scaling optical communications for the next decade and beyond," *Bell Labs Technical Journal*, vol. 14, no. 4, pp. 3–9, 2010.

[11] E. Modone and G. Roba, "Oh reduction in preforms by isotope exchange," *Electronics Letters*, vol. 17, no. 21, pp. 815–817, 1981.

[12] D. Richardson, J. Fini, and L. E. Nelson, "Space-division multiplexing in optical fibres," *Nature Photonics*, vol. 7, no. 5, p. 354, 2013.

[13] G. Kramer, M. I. Yousefi, and F. R. Kschischang, "Upper bound on the capacity of a cascade of nonlinear and noisy channels," in *2015 IEEE Information Theory Workshop (ITW)*, IEEE, 2015, pp. 1–4.

[14] T. Schneider, *Nonlinear optics in telecommunications*. Springer Science & Business Media, 2004.

[15] A. Shabat and V. Zakharov, "Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media," *Soviet physics JETP*, vol. 34, no. 1, p. 62, 1972.

[16] A. Hasegawa and F. Tappert, "Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers. i. anomalous dispersion," *Applied Physics Letters*, vol. 23, no. 3, pp. 142–144, 1973.

[17] A. Hasegawa and F. Tappert, "Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers. ii. normal dispersion," *Applied Physics Letters*, vol. 23, no. 4, pp. 171–172, 1973.

[18] P. D. Lax, "Integrals of nonlinear equations of evolution and solitary waves," *Communications on pure and applied mathematics*, vol. 21, no. 5, pp. 467–490, 1968.

[19] A. Hasegawa and T. Nyu, "Eigenvalue communication," *Journal of lightwave technology*, vol. 11, no. 3, pp. 395–399, 1993.

[20] E. Meron, M. Feder, and M. Shtaif, "On the achievable communication rates of generalized soliton transmission systems," *arXiv preprint arXiv:1207.0297*, 2012.

[21] 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*, vol. 113, no. 1, p. 013 901, 2014.

[22] S. Hari, F. Kschischang, and M. Yousefi, "Multi-eigenvalue communication via the nonlinear fourier transform," in *2014 27th Biennial Symposium on Communications (QBSC)*, IEEE, 2014, pp. 92–95.

[23] M. Yousefi and X. Yangzhang, "Linear and nonlinear frequency-division multiplexing," *IEEE Transactions on Information Theory*, 2019.

[24] X. Yangzhang, M. I. Yousefi, A. Alvarado, D. Lavery, and P. Bayvel, "Nonlinear frequency-division multiplexing in the focusing regime," in *Optical Fiber Communication Conference*, Optical Society of America, 2017, Tu3D–1.

[25] H. Bülow, "Experimental demonstration of optical signal detection using nonlinear fourier transform," *Journal of Lightwave Technology*, vol. 33, no. 7, pp. 1433–1439, 2015.

[26] A. Geisler and C. G. Schaeffer, "Experimental nonlinear frequency division multiplexed transmission using eigenvalues with symmetric real part," in *ECOC 2016; 42nd European Conference on Optical Communication*, VDE, 2016, pp. 1–3.

[27] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear fourier transform, part ii: Numerical methods," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 4329–4345, 2014.

[28] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear fourier transform, part iii: Spectrum modulation," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 4346–4369, 2014.

[29] S. K. Turitsyn, J. E. Prilepsky, S. T. Le, *et al.*, "Nonlinear fourier transform for optical data processing and transmission: Advances and perspectives," *Optica*, vol. 4, no. 3, pp. 307–322, 2017.

[30] V. Vaibhav and S. Wahls, "Introducing the fast inverse nft," in *2017 Optical Fiber Communications Conference and Exhibition (OFC)*, IEEE, 2017, pp. 1–3.

[31] S. Wahls and H. V. Poor, "Introducing the fast nonlinear fourier transform," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2013, pp. 5780–5784.

[32] S. Civelli, E. Forestieri, and M. Secondini, "Why noise and dispersion may seriously hamper nonlinear frequency-division multiplexing," *IEEE Photonics Technology Letters*, vol. 29, no. 16, pp. 1332–1335, 2017.

[33] V. Aref, S. T. Le, and H. Buelow, "Does the cross-talk between nonlinear modes limit the performance of nfdm systems?" In *2017 European Conference on Optical Communication (ECOC)*, IEEE, 2017, pp. 1–3.

[34] I. N. Bronstein, J. Hromkovic, B. Luderer, *et al.*, *Taschenbuch der mathematik*. Springer-Verlag, 2012, vol. 1.

[35] J. G. Proakis and M. Salehi, *Digital communications*. McGraw-hill New York, 2001, vol. 4.

[36] R. B. Ash, *Basic probability theory*. Courier Corporation, 2008.

[37] C. M. Grinstead and J. L. Snell, *Introduction to probability*. American Mathematical Soc., 2012.

[38] C. E. Shannon and W. Weaver, *The mathematical theory of communication*. University of Illinois press, 1998.

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

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

[41] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[42] M. I. Yousefi, G. Kramer, and F. R. Kschischang, "Upper bound on the capacity of the nonlinear schrödinger channel," in *2015 IEEE 14th Canadian Workshop on Information Theory (CWIT)*, IEEE, 2015, pp. 22–26.

[43] M. I. Yousefi and F. R. Kschischang, "Information transmission using the nonlinear fourier transform, part i: Mathematical tools," *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 4312–4328, 2014.

[44] I. Gohberg and S. Goldberg, *Basic operator theory*. Birkhäuser, 2013.

[45] C. S. Kubrusly, *Elements of operator theory*. Springer, 2011.

[46] T. Kato, *Perturbation theory for linear operators*. Springer Science & Business Media, 2013, vol. 132.

[47] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn, "Nonlinear inverse synthesis technique for optical links with lumped amplification," *Optics express*, vol. 23, no. 7, pp. 8317–8328, 2015.

[48] F. J. Garcia-Gomez, "Numerically computing achievable rates of memoryless channels," *TUM University Library*, 2019. [Online]. Available: `https://mediatum.ub.tum.de/node?id=1533663`.

[49] G. Kaplan and S. Shamai, "Information rates and error exponents of compound channels with application to antipodal signaling in a fading environment," *AEU. Archiv für Elektronik und Übertragungstechnik*, vol. 47, no. 4, pp. 228–239, 1993.

[50] G. P. Agrawal, *Fiber-optic communication systems*. John Wiley & Sons, 2012.

[51] G. P. Agrawal, "Nonlinear fiber optics academic," *New York*, vol. 19892, p. 44, 2013.

[52] T. Fehenberger, "Analysis and optimization of coded modulation for nonlinear fiber-optic communication systems," Ph.D. dissertation, Technische Universität München, 2017.

[53] G. P. Agrawal, "Nonlinear fiber optics: Its history and recent progress," *JOSA B*, vol. 28, no. 12, A1–A10, 2011.

[54] N. Hanik, *Optical communication systems (lecture notes)*, 2020.

[55] C. Poole and R. Wagner, "Phenomenological approach to polarisation dispersion in long single-mode fibres," *Electronics Letters*, vol. 19, no. 22, pp. 1029–1030, 1986.

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

[57] O. V. Sinkin, R. Holzlöhner, J. Zweck, and C. R. Menyuk, "Optimization of the split-step fourier method in modeling optical-fiber communications systems," *Journal of lightwave technology*, vol. 21, no. 1, p. 61, 2003.

[58] C. Häger and H. D. Pfister, "Deep learning of the nonlinear schrödinger equation in fiber-optic communications," in *2018 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2018, pp. 1590–1594.

[59] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex fourier series," *Mathematics of computation*, vol. 19, no. 90, pp. 297–301, 1965.

[60] G. H. Weiss and A. A. Maradudin, "The baker-hausdorff formula and a problem in crystal physics," *Journal of Mathematical Physics*, vol. 3, no. 4, pp. 771–777, 1962.

[61] 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*, vol. 31, no. 3, pp. 398–406, 2012.

[62] A. Mecozzi, C. Antonelli, and M. Shtaif, "Nonlinear propagation in multi-mode fibers in the strong coupling regime," *Optics express*, vol. 20, no. 11, pp. 11 673–11 678, 2012.

[63] K.-P. Ho and J. M. Kahn, "Statistics of group delays in multimode fiber with strong mode coupling," *Journal of lightwave technology*, vol. 29, no. 21, pp. 3119–3128, 2011.

[64] K.-P. Ho and J. M. Kahn, "Mode-dependent loss and gain: Statistics and effect on mode-division multiplexing," *Optics express*, vol. 19, no. 17, pp. 16 612–16 635, 2011.

[65] A. Mecozzi and R.-J. Essiambre, "Nonlinear shannon limit in pseudolinear coherent systems," *Journal of Lightwave Technology*, vol. 30, no. 12, pp. 2011–2024, 2012.

[66] C. Antonelli, M. Shtaif, and A. Mecozzi, "Modeling of nonlinear propagation in space-division multiplexed fiber-optic transmission," *Journal of Lightwave Technology*, vol. 34, no. 1, pp. 36–54, 2015.

[67] C. Headley and G. P. Agrawal, *Raman amplification in fiber optical communication systems*. Academic press, 2005.

[68] L. D. Coelho, "Modeling, simulation and optimization of optical communication systems using advanced modulation formats," Dissertation, Technische Universität München, München, 2010.

[69] Q. Lin and G. P. Agrawal, "Raman response function for silica fibers," *Optics letters*, vol. 31, no. 21, pp. 3086–3088, 2006.

[70] A. H. Beshr and M. H. Aly, "Noise figure of distributed raman amplifier with different pumping configurations in s-band: A new approach," *Alexandria Engineering Journal*, 2020.

[71] J. Bromage, "Raman amplification for fiber communications systems," *journal of lightwave technology*, vol. 22, no. 1, p. 79, 2004.

[72] N. J. Muga, M. C. Fugihara, M. F. S. Ferreira, and A. N. Pinto, "Non-gaussian ase noise in raman amplification systems," *Journal of Lightwave Technology*, vol. 27, no. 16, pp. 3389–3398, 2009.

[73] A. Kobyakov, M. Vasilyev, S. Tsuda, G. Giudice, and S. Ten, "Analytical model for raman noise figure in dispersion-managed fibers," *IEEE Photonics Technology Letters*, vol. 15, no. 1, pp. 30–32, 2003.

[74] P. Hansen, L. Eskildsen, J. Stentz, *et al.*, "Rayleigh scattering limitations in distributed raman pre-amplifiers," *IEEE Photonics Technology Letters*, vol. 10, no. 1, pp. 159–161, 1998.

[75] N. J. Muga, M. C. Fugihara, M. F. Ferreira, and A. N. Pinto, "Ase noise simulation in raman amplification systems," in *Conf. on Telecommunications-ConfTele*, Citeseer, 2009.

[76] M. Abdelaziz and T. A. Gulliver, "Triangular constellations for adaptive modulation," *IEEE Transactions on Communications*, vol. 66, no. 2, pp. 756–766, 2017.

[77] N. S. Bergano, "Wavelength division multiplexing in long-haul transoceanic transmission systems," *Journal of lightwave technology*, vol. 23, no. 12, pp. 4125–4139, 2005.

[78] D. Chadha, *Optical WDM Networks: From Static to Elastic Networks*. John Wiley & Sons, 2019.

[79] N. K. Fontaine, "Spectrally-sliced coherent receivers for thz bandwidth optical communications," in *39th European Conference and Exhibition on Optical Communication (ECOC 2013)*, IET, 2013, pp. 1–3.

[80] S. Daumont, B. Rihawi, and Y. Lout, "Root-raised cosine filter influences on papr distribution of single carrier signals," in *2008 3rd International Symposium on Communications, Control and Signal Processing*, IEEE, 2008, pp. 841–845.

[81] J. Armstrong, "Ofdm for optical communications," *Journal of lightwave technology*, vol. 27, no. 3, pp. 189–204, 2009.

[82] W. Shieh and I. Djordjevic, *OFDM for optical communications*. Academic press, 2009.

[83] J. G. Proakis, M. Salehi, N. Zhou, and X. Li, *Communication systems engineering*. Prentice Hall New Jersey, 1994, vol. 2.

[84] G. S. He, "Optical phase conjugation: Principles, techniques, and applications," *Progress in Quantum Electronics*, vol. 26, no. 3, pp. 131–191, 2002.

[85] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *Journal of Lightwave Technology*, vol. 26, no. 20, pp. 3416–3425, 2008.

[86] A. Napoli, Z. Maalej, V. A. Sleiffer, *et al.*, "Reduced complexity digital backpropagation methods for optical communication systems," *Journal of lightwave technology*, vol. 32, no. 7, pp. 1351–1362, 2014.

[87] C. Häger and H. D. Pfister, "Nonlinear interference mitigation via deep neural networks," in *Optical fiber communication conference*, Optical Society of America, 2018, W3A–4.

[88] G. Liga, T. Xu, A. Alvarado, R. I. Killey, and P. Bayvel, "On the performance of multichannel digital backpropagation in high-capacity long-haul optical transmission," *Optics Express*, vol. 22, no. 24, pp. 30 053–30 062, 2014.

[89] M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, "The inverse scattering transform-fourier analysis for nonlinear problems," *Studies in Applied Mathematics*, vol. 53, no. 4, pp. 249–315, 1974.

[90] M. J. Ablowitz and H. Segur, *Solitons and the inverse scattering transform*. Siam, 1981, vol. 4.

[91] P. Kazakopoulos and A. L. Moustakas, "Nonlinear schrödinger equation with random gaussian input: Distribution of inverse scattering data and eigenvalues," *Physical Review E*, vol. 78, no. 1, p. 016 603, 2008.

[92] T. Trogdon and S. Olver, "Numerical inverse scattering for the focusing and defocusing nonlinear schrödinger equations," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 469, no. 2149, p. 20 120 330, 2013.

[93] R. Dodd, "Ld faddeev and la takhtajan, hamiltonian methods in the theory of solitons," *Bulletin (New Series) of the American Mathematical Society*, vol. 19, no. 2, pp. 565–568, 1988.

[94] M. Kamalian, J. E. Prilepsky, S. T. Le, and S. K. Turitsyn, "Spectral efficiency estimation in periodic nonlinear fourier transform based communication systems," in *2017 Optical Fiber Communications Conference and Exhibition (OFC)*, IEEE, 2017, pp. 1–3.

[95] J.-W. Goossens, H. Hafermann, and Y. Jaouën, "Data transmission based on exact inverse periodic nonlinear fourier transform, part i: Theory," *Journal of Lightwave Technology*, vol. 38, no. 23, pp. 6499–6519, 2020.

[96] M. Beck, G. Marchesi, D. Pixton, and L. Sabalka, *A first course in complex analysis*. Orthogonal Publishing, 2018.

[97] F. J. García-Gómez, "Communication using eigenvalues of higher multiplicity of the nonlinear fourier transform," *Journal of Lightwave Technology*, vol. 36, no. 23, pp. 5442–5450, 2018.

[98] H. Flaschka, A. Newell, and M. Tabor, "Integrability," in *What is integrability?* Springer, 1991, pp. 73–114.

[99] V. E. Zakharov *et al.*, *What is integrability?* Springer, 1991.

[100] V. E. Zakharov and L. D. Faddeev, "Korteweg–de vries equation: A completely integrable hamiltonian system," *Funktsional'nyi Analiz i ego Prilozheniya*, vol. 5, no. 4, pp. 18–27, 1971.

[101] V. Vaibhav, "Higher order convergent fast nonlinear fourier transform," *IEEE Photonics Technology Letters*, vol. 30, no. 8, pp. 700–703, 2018.

[102] S. Chimmalgi, P. J. Prins, and S. Wahls, "Fast nonlinear fourier transform algorithms using higher order exponential integrators," *IEEE Access*, vol. 7, pp. 145 161–145 176, 2019.

[103] G. P. Agrawal, "Nonlinear fiber optics," in *Nonlinear Science at the Dawn of the 21st Century*, Springer, 2000, pp. 195–211.

[104] J. Satsuma and N. Yajima, "B. initial value problems of one-dimensional self-modulation of nonlinear waves in dispersive media," *Progress of Theoretical Physics Supplement*, vol. 55, pp. 284–306, 1974.

[105] P. J. Davis, "Leonhard euler's integral: A historical profile of the gamma function: In memoriam: Milton abramowitz," *The American Mathematical Monthly*, vol. 66, no. 10, pp. 849–869, 1959.

[106] V. Aref, "Control and detection of discrete spectral amplitudes in nonlinear fourier spectrum," *arXiv preprint arXiv:1605.06328*, 2016.

[107] K. E. Atkinson, *An introduction to numerical analysis*. John Wiley & Sons, 2008.

[108] R. Bulirsch, J. Stoer, and J. Stoer, *Introduction to numerical analysis*. Springer, 2002, vol. 3.

[109] M. Ablowitz and J. Ladik, "Nonlinear differential–difference equations and fourier analysis," *Journal of Mathematical Physics*, vol. 17, no. 6, pp. 1011–1018, 1976.

[110] T. Tao and C. Thiele, "Nonlinear fourier analysis," *arXiv preprint arXiv:1201.5129*, 2012.

[111] I. S. Chekhovskoy, S. Medvedev, I. Vaseva, E. Sedov, and M. P. Fedoruk, "Introducing phase jump tracking-a fast method for eigenvalue evaluation of the direct zakharov-shabat problem," *Communications in Nonlinear Science and Numerical Simulation*, p. 105 718, 2020.

[112] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3d surface construction algorithm," *ACM siggraph computer graphics*, vol. 21, no. 4, pp. 163–169, 1987.

[113] A. Span, V. Aref, H. Bülow, and S. ten Brink, "Successive eigenvalue removal for multi-soliton spectral amplitude estimation," *Journal of Lightwave Technology*, vol. 38, no. 17, pp. 4708–4714, 2020.

[114] N. H. Benedikt Leible Daniel Plabst, "Back-to-back performance of the full spectrum nonlinear fourier transform and its inverse," en, *Entropy*, vol. 22, no. 10, p. 1131, 2020. DOI: `https://doi.org/10.3390/e22101131`.

[115] N. H. Benedikt Leible Daniel Plabst, "Stability of the full spectrum nonlinear fouriertransform," in *International Conference on Transparent Optical Networks (ICTON) 2020*, Jul. 2020.

[116] A. Span, V. Aref, H. Bülow, and S. Ten Brink, "On time-bandwidth product of multi-soliton pulses," in *2017 IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2017, pp. 61–65.

[117] F. J. García-Gómez, "Achievable rates of nonlinear fourier transform-based optical communication systems," en, in *Workshop Nichtlineare Fourier Transformation der Christian-Albrechts-Universität zu Kiel*, Christian-Albrechts-Universität zu Kiel, 2017.

[118] B. Leible, Y. Chen, M. I. Yousefi, and N. Hanik, "Soliton transmission with 5 eigenvalues over 2000km of raman-amplified fiber," in *2018 20th International Conference on Transparent Optical Networks (ICTON)*, IEEE, 2018, pp. 1–4.

[119] F. Ahmad and M. Razzaghi, "A numerical solution to the gel'fand-levitan-marchenko equation," *Applied mathematics and computation*, vol. 89, no. 1-3, pp. 31–39, 1998.

[120] L. L. Frumin, A. Gelash, and S. K. Turitsyn, "New approaches to coding information using inverse scattering transform," *Physical Review Letters*, vol. 118, no. 22, p. 223 901, 2017.

[121] S. Wahls and H. V. Poor, "Inverse nonlinear fourier transforms via interpolation: The ablowitz-ladik case," in *Proceeding of International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, Citeseer, 2014, pp. 1848–1855.

[122] O. Wright, "The darboux transformation of some manakov systems," *Applied mathematics letters*, vol. 16, no. 5, pp. 647–652, 2003.

[123] V. Aref, S. T. Le, and H. Buelow, "Demonstration of fully nonlinear spectrum modulated system in the highly nonlinear optical transmission regime," in *ECOC 2016-Post Deadline Paper; 42nd European Conference on Optical Communication*, VDE, 2016, pp. 1–3.

[124] V. Aref, S. T. Le, and H. Buelow, "Modulation over nonlinear fourier spectrum: Continuous and discrete spectrum," *Journal of Lightwave Technology*, vol. 36, no. 6, pp. 1289–1295, 2018.

[125] O. Belai, L. Frumin, E. Podivilov, and D. Shapiro, "Efficient numerical method of the fiber bragg grating synthesis," *JOSA B*, vol. 24, no. 7, pp. 1451–1457, 2007.

[126] L. L. Frumin, O. V. Belai, E. V. Podivilov, and D. A. Shapiro, "Efficient numerical method for solving the direct zakharov–shabat scattering problem," *JOSA B*, vol. 32, no. 2, pp. 290–296, 2015.

[127] S. Medvedev, I. Vaseva, I. Chekhovskoy, and M. Fedoruk, "Exponential fourth order schemes for direct zakharov-shabat problem," *Optics Express*, vol. 28, no. 1, pp. 20–39, 2020.

[128] A. Geisler, "Experimental investigation of information transmission using the nonlinear fourier transformation," Ph.D. dissertation, Universitätsbibliothek der HSU/UniBwH, 2022.

[129] J. Garcia and V. Aref, "Statistics of the eigenvalues of a noisy multi-soliton pulse," in *2018 European Conference on Optical Communication (ECOC)*, IEEE, 2018, pp. 1–3.

[130] F. J. García-Gómez and V. Aref, "Statistics of the nonlinear discrete spectrum of a noisy pulse," *Journal of Lightwave Technology*, vol. 37, no. 14, pp. 3563–3570, 2019.

[131] S. T. Le, J. E. Prilepsky, P. Rosa, J. D. Ania-Castañón, and S. K. Turitsyn, "Nonlinear inverse synthesis for optical links with distributed raman amplification," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1778–1786, 2015.

[132] P. B. de Koster and S. Wahls, "Fibre model identification for nonlinear fourier transform-based transmission," 2019.

[133] P. De Koster and S. Wahls, "Dispersion and nonlinearity identification for single-mode fibers using the nonlinear fourier transform," *Journal of Lightwave Technology*, vol. 38, no. 12, pp. 3252–3260, 2020.

[134] P. De Koster, J. Koch, O. Schulz, S. Pachnicke, and S. Wahls, "Experimental validation of nonlinear fourier transform-based kerr-nonlinearity identification over a 1600 km ssmf link," in *Optical Fiber Communication Conference*, Optica Publishing Group, 2022, W2A–39.

[135] B. Leible and N. Hanik, "Eigenvalue trajectories in multispan soliton transmission systems under lumped and distributed amplification," in *2019 21st International Conference on Transparent Optical Networks (ICTON)*, IEEE, 2019, pp. 1–4.

[136] G. Frank, "Pulse code communication," *US Patent A*, vol. 2632058, 1953.

[137] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[138] D. Zibar, O. Winther, N. Franceschi, *et al.*, "Nonlinear Impairment Compensation Using Expectation Maximization for Dispersion Managed and Unmanaged pdm 16-qam Transmission," *optex*, vol. 20, no. 26, B181–B196, 2012.

[139] H. W. Kuhn, "The hungarian method for the assignment problem," *Naval research logistics quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.

[140] H. Buelow, V. Aref, and W. Idler, "Transmission of waveforms determined by 7 eigenvalues with psk-modulated spectral amplitudes," in *ECOC 2016; 42nd European Conference on Optical Communication*, VDE, 2016, pp. 1–3.

[141] J.-W. Goossens, M. I. Yousefi, Y. Jaouën, and H. Hafermann, "Polarization-division multiplexing based on the nonlinear fourier transform," *Optics express*, vol. 25, no. 22, pp. 26 437–26 452, 2017.

[142] 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 *2017 European Conference on Optical Communication (ECOC)*, IEEE, 2017, pp. 1–3.

[143] 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*, vol. 5, no. 3, pp. 263–270, 2018.

[144] W. A. Gemechu, T. Gui, J.-W. Goossens, *et al.*, "Dual polarization nonlinear frequency division multiplexing transmission," *IEEE Photonics Technology Letters*, vol. 30, no. 18, pp. 1589–1592, 2018.

[145] S. V. Manakov, "On the theory of two-dimensional stationary self-focusing of electromagnetic waves," *Soviet Physics-JETP*, vol. 38, no. 2, pp. 248–253, 1974.

[146] A. Degasperis and S. Lombardo, "Integrability in action: Solitons, instability and rogue waves," in *Rogue and shock waves in nonlinear dispersive media*, Springer, 2016, pp. 23–53.

[147] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: Analysis and implementation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 881–892, 2002.