

**Some pages of this thesis may have been removed for copyright restrictions.**

If you have discovered material in Aston Research Explorer which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our [Takedown policy](#) and contact the service immediately (openaccess@aston.ac.uk)

# NONLINEAR FOURIER TRANSFORM IN APPLICATION TO LONG-HAUL OPTICAL COMMUNICATIONS

Anastasiia Vasylchenkova  
Doctor of Philosophy



Aston Institute of Photonic Technologies  
ASTON UNIVERSITY

© Anastasiia Vasylchenkova, 2020

Anastasiia Vasylchenkova asserts her moral right to be identified as the author of this  
thesis

This copy of the thesis has been supplied on condition that anyone who consults it is  
understood to recognise that its copyright rests with its author and that no quotation  
from the thesis and no information derived from it may be published without  
appropriate permission or acknowledgement.

June 2020



# Declaration

I declare that I have produced this work without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This work has not previously been presented in identical or similar form to any degree examination board.



# Nonlinear Fourier Transform In Application To Long-Haul Optical Communications

Anastasiia Vasylychenkova

Doctor of Philosophy, June 2020

The optical fibres form a basis of the long-haul transmissions systems, and is a significant component of the connectivity infrastructure. Rapidly growing demand in data traffic requires instantaneous imperative actions with long-term effect to meet the future expansion of the digital economy. The current optical networks resources are overstretched, and the further extensive utilisation will ultimately constrain the development of other economic sectors.

The intelligent and effective usage of the installed infrastructure can shift forward the existent limitations, keeping the cost low because of avoiding of the reinstallation. One of the principal constrain, which bounds the further optical fibre capacity grows, is the existence of undesirable nonlinear phenomena, the so-called Kerr nonlinearity, causing self-phase, cross-phase modulation and four-wave mixing. The combination of advanced achievements of mathematical physics, together with communication engineering and information theory allowed to implement the so-called nonlinear Fourier transform (NFT) approach to optical communication. In its paradigm, the fibre nonlinearity is considered as a valuable part of the model, and the NFT mapping effectively (de)composes the signal to naturally non-interacting modes. The NFT concept can be applied to the signal propagation model with either vanishing or periodic boundary condition, which involves the different structures of parameters for manipulation.

In this thesis, I focused on the investigation of boundary condition cases, discovering analytical properties, available degrees of freedom, developing numerical methods, and coding approaches; then examining their performance via the simulation of optical transmission systems. The results allow us to conclude the existence of several technical limitations, which limit the achievable transmission quality and data-rate. These include: the deviation of the channel model from the purely integrable, nonlinear and not explicit coupling of the resulting signal parameters, numerical methods accuracy and amplifiers noise accumulation. In spite of those, the simulations demonstrate the considerable performance of NFT-based communication systems.

Keywords: optical communication, nonlinearity mitigation, signal processing



# Acknowledgements

That happened that there were four people in total, who supervised my PhD research. They gave me a great boost of motivation and managing their not very coherent tasks and delegations, so in the first instance, I would like to express my gratitude to them. I would like to thank Prof Sergei K. Turitsyn, who despite his overwhelmed schedule with duties as a Director of Institute and as an active investigator of plenty of research projects, managed to firmly direct me in my PhD journey. He allowed me a great deal of research freedom and opportunity to find my own path in the scientific community, actively encouraging my development in management, leadership and raising my confidence. I also would like to thank my co-supervisor Dr Yaroslav Prylepskiy, who dedicated a great amount of his time in teaching me details of research work, paper publications, and task managing. The last but not least of my co-supervisors, Dr Amit Chattopadhyay and Dr Sotos Generalis expressed their interest and support to my work and development in a really encouraging way, what is highly appreciated.

I cannot leave unmentioned my colleagues, without whom this work would not be such good as it is and who noticeably contributed to the obtained results. I am grateful to my senior colleague Dr Morteza Kamalian Kopae, who despite his involvement in the various project at Aston Institute of Photonic Technologies (AiPT) was kind and supportive, guiding me through the details of NFT numerical routines, transmission simulations and supervisor management. I would like to acknowledge Prof Dmitry Shepelsky from Kharkiv National University, for his contribution to the development of the novel methods in optical communications, many of good research publications, done together, and for being a great role model of a researcher in mathematical physics. The most detached person among my close collaborators, Dr Stanislav Derevyanko from Ben-Gurion University, acted for me as an occasional mentor and served me as a long-term role and career model of a successful and well-performing theoretical physicist. The remaining members of the NFT research group at AiPT, Dr Maryna Pankratova and Dr Oleksandr Kotlyar, were inestimably helpful in raising my confidence as a researcher and mentor. I am glad that we developed firm liaisons of friendship and solidarity. I would also thank Dr Nikolay Chichkov, a close collaborator from another research group at AiPT and a co-author of several research publications, for his continuous career advice and many warm and cherished discussions about science, research and beyond.

Apart from the researchers, our institute is full of bright and smart people, who always inspired me with their professionalism and kindness. I am especially thankful to Ms Tatiana Kilina, who provided me with remarkably valuable mentoring on the way of my growth and maturation professionally and beyond. I am also grateful to the entire cosy and kind AiPT community, where I spent all the time of my PhD study. The friendly and supportive environment in the institute helped me to overcome the barriers in my study, research, personal and professional development, and especially in socialising and accommodating to a new country. The special thank to my PhD-students-office mates, who were extremely patient to my stopless nerves and tears in the last weeks before this thesis submission, overlapped with the individual fellowship proposal submission, job-search frustration and personal life problems.

Beyond my group and institute colleagues, I cannot leave unmentioned the valuable and priceless support from my dad and my friends, who did not stop reminding me that I was smart, talented, strong and loved, and for their inspirational environment within my school friends community, the LOL Summer Camp and the Young/Student Physicists' Tournaments, which elevated my belief in humanity.

# Contents

<b>1</b>	<b>Introduction</b>	<b>19</b>
1.1	Motivation . . . . .	19
1.2	Thesis outline . . . . .	19
1.3	Contribution . . . . .	20
1.4	Tools . . . . .	20
<b>2</b>	<b>Current state of art</b>	<b>22</b>
2.1	Physical effects in the optical fibre . . . . .	22
2.1.1	General remarks . . . . .	22
2.1.2	Group-velocity dispersion . . . . .	22
2.1.3	Nonlinear effects . . . . .	23
2.1.4	Loss and amplification . . . . .	23
2.2	Optical fibre modelling with Nonlinear Schrodinger equation . . . . .	25
2.2.1	Path averaging . . . . .	26
2.2.2	Normalisation approaches . . . . .	26
2.3	Numerical solution of initial value problem for NLS . . . . .	27
2.4	Optical communications . . . . .	28
2.5	Nonlinearity mitigation techniques . . . . .	29
2.6	Additional details for transmission simulations . . . . .	30
2.6.1	Modulation approaches . . . . .	31
2.6.2	Single-mode fibre parameters . . . . .	31
2.6.3	Transmission quality metrics . . . . .	31
2.6.4	Noise consideration in the numerical simulations . . . . .	33
<b>3</b>	<b>Inverse scattering transform</b>	<b>35</b>
3.1	Vanishing boundary conditions . . . . .	38
3.1.1	Local definitions . . . . .	38
3.1.2	Alternative determination of the scattering data via ZSP for enveloping functions . . . . .	41
3.1.3	Relations between soliton parameters and spectral data . . . . .	42
3.1.4	Evolution of scattering data . . . . .	43
3.1.5	Discussion . . . . .	44
3.2	Periodic boundary conditions . . . . .	44
3.2.1	Local definitions . . . . .	44
3.2.2	Evolution of scattering data . . . . .	47
3.2.3	Discussion . . . . .	48
3.3	Inverse NFT . . . . .	48
3.3.1	General remarks . . . . .	48
3.3.2	Gelfand-Levitan-Marchenko equations . . . . .	50
3.3.3	Darboux transform . . . . .	51
3.3.4	RHP for vanishing boundary condition . . . . .	52

3.3.5	RHP in finite-band case . . . . .	56
3.3.6	Alternative formulation of RHP in the finite-band case . . . . .	59
3.3.7	Algebro-geometric approach . . . . .	60
3.3.8	Inverse problem and signal generation discussion . . . . .	62
<b>4</b>	<b>Numerical methods for scattering data evaluation</b>	<b>64</b>
4.1	Test profiles and metrics . . . . .	64
4.2	Transfer-matrix methods . . . . .	66
4.2.1	Derivative computation . . . . .	72
4.3	Iterative methods for discrete eigenvalues . . . . .	73
4.4	Fourier collocation method . . . . .	78
4.5	Discrete spectrum methods discussion . . . . .	80
<b>5</b>	<b>Contour integrals within inverse scattering problem</b>	<b>83</b>
5.1	Contour integrals for zero location . . . . .	83
5.1.1	Mathematical details . . . . .	83
5.1.2	Numerical accuracy of the contour integrals for eigenvalues location . . . . .	85
5.2	Multisolitonic test of the discrete eigenvalues methods . . . . .	87
5.3	The contour integrals method paired with iterative search . . . . .	87
5.4	Accuracy and channel noise tolerance . . . . .	89
5.4.1	Discretisation and additive noise influence . . . . .	89
5.4.2	Closely located eigenvalues . . . . .	93
5.4.3	Influence of noise in the NLS channel . . . . .	93
5.5	Contour integrals approach discussion . . . . .	97
<b>6</b>	<b>Optical communication schemes based on NFT</b>	<b>98</b>
6.1	Overview . . . . .	98
6.2	Conventional NFT transmission . . . . .	98
6.2.1	$r$ -modulation . . . . .	99
6.3	Contour integrals communication . . . . .	102
6.4	$b$ -modulation NFT transmission . . . . .	104
6.4.1	Formal justification of signal and auxiliary $b$ -function correspondence . . . . .	105
6.4.2	Localised $b$ -modulated profiles containing solitons . . . . .	112
6.4.3	$b$ -modulation transmission systems . . . . .	117
6.5	Periodic boundary conditions . . . . .	120
6.5.1	Signal properties for the RHP-generated profiles . . . . .	122
6.5.2	Numerical accuracy of RHP-based processing . . . . .	124
6.5.3	Main spectrum modulation in RHP-based transmission . . . . .	125
6.5.4	Phases modulation in RHP-based transmission . . . . .	130
6.6	Communication systems discussion . . . . .	134
<b>7</b>	<b>Overall discussion and conclusion</b>	<b>136</b>
<b>A</b>	<b>Fourier transform in MATLAB</b>	<b>149</b>
<b>B</b>	<b>Topical visuals</b>	<b>151</b>

# List of Tables

2.1	Listing of the split-step coefficients for different forms of optical fibre models based on NLS. . . . .	28
2.2	Conventional modulation constellation structures. . . . .	32
2.3	Single-mode fibre parameters in the vicinity of C-band. . . . .	33
4.1	Runtimes of continuous energy evaluation for different NFT algorithms. . .	72
5.1	Runtimes of eigenvalues evaluation for different contour shapes. . . . .	86
5.2	Normalised runtimes of eigenvalues evaluation for iterative and contour integrals methods. . . . .	89
6.1	The values of $a_m$ for the flat-top waveform, eq. (6.45) (cited from [47]). . .	113

# List of Figures

2.1	The attenuation coefficient of silica fibre near the transparency point as a function of the incident wave frequency (cited from [122]). . . . .	24
2.2	The achievable data capacity by optical fibre communication systems, as a function of link length and SNR (cited from [36]). . . . .	29
3.1	Schematic of the procedure of integrable equation solution using the IST approach. . . . .	36
3.2	Visual representation of the NF image for the vanishing boundary condition case. . . . .	40
3.3	Visual representation of the NF image for the finite-band case. . . . .	45
3.4	The sech potential reconstruction via RHP for vanishing boundary condition via different jump matrices expressions. . . . .	54
3.5	The contours $\Gamma$ , used for the RHP in the vanishing boundary condition case for nonzero $\zeta$ to deal with jump matrices convergence. . . . .	55
3.6	The numerical solution of RHP in the vanishing boundary condition case for $\zeta = 10$ for two contours deformations. . . . .	56
3.7	Closed contours $\Xi_{0,1}$ , enveloping the arcs $\Gamma_{0,1}$ for alternative RHP formulation, which avoids singularity points. . . . .	60
4.1	Transparency scale for amplitude variation in grey tones; progressively darker shades indicate higher amplitudes . . . . .	68
4.2	MSRE for the computation of the NF scattering functions as a function of the number of discretization points and signal amplitude. . . . .	69
4.3	Relative error of the continuous spectrum energy evaluation as a function of a number of discretization points and signal amplitude. . . . .	70
4.4	Relative error of the NF scattering functions as a function of spectral parameter and signal amplitude. . . . .	71
4.5	Relative error for $a'(\lambda)$ computation as a function of the spectral parameter and the number of discretisation points. . . . .	74
4.6	The regions of initial assumptions for the zero approximation in which the iterative algorithms reach correct zero with less than 1% relative error for soliton potential. . . . .	76
4.7	The regions of initial assumptions for the zero approximation in which the iterative algorithms reach correct zero with less than 1% relative error for rectangular potential. . . . .	77
4.8	Relative error and normalised computational time of discrete eigenvalue computed using two types of FC method in dependence on number of points $n$ . . . . .	81
5.1	Relative error as a function on the number of contour discretization points for different contour shapes and contour integrals methods. . . . .	86

5.2	Relative error of the eigenvalue location via root-finding iterative and contour integrals methods for multieigenvalue profile with computational runtime.	88
5.3	Relative error of the eigenvalue location via the hybrid method for multieigenvalue profile with computational runtime.	89
5.4	EVM <sup>2</sup> as a function of time and NF domain sampling for different number of eigenvalues.	90
5.5	EVM <sup>2</sup> of the potential with AWGN for different eigenvalues constellations.	91
5.6	EVM <sup>2</sup> for degenerate eigenvalue evaluation using degenerate breather profile in dependence on the time domain discretisation.	92
5.7	Deviation of numerically evaluated eigenvalues of the noiseless degenerate breather profile computed via contour integrals and iterative method.	92
5.8	EVM <sup>2</sup> of the degenerate breather with AWGN eigenvalue evaluation.	93
5.9	Relative error of numerically evaluated eigenvalues by contour integrals and iterative algorithms for closely located eigenvalues.	94
5.10	Normalised variance of the eigenvalues and contour integrals for multieigenvalue profiles after the propagation through NLS channel with AWGN.	96
6.1	General concept of NFT-based communication.	99
6.2	The principal scheme of $r$ -modulation transmission system.	99
6.3	$Q^2$ factor vs. power for $r$ -modulation in the noiseless channel.	100
6.4	$Q^2$ factor vs. power for $r$ -modulation in the presence of ASE noise.	100
6.5	The optimal $Q^2$ factor for different propagation distances in $r$ -modulation.	101
6.6	The processing time required to reach the maximum $Q^2$ at the Tx and Rx for different transmission distances $r$ -modulation.	101
6.7	Schematics of one-to-one correspondence between NLS profile, eigenvalues and contour integrals.	102
6.8	The schemes of (a) the multieigenvalue communication concept and (b) the contour integrals communication concept.	103
6.9	Regions of values of $s_p$ , $\lambda_j$ and $\hat{\lambda}_j$ in the corresponding complex planes for comparison of multieigenvalue and contour integrals communication.	103
6.10	Illustration of one-to-one correspondence between bandlimited function $b$ and the corresponding time-domain profile, highlighting the contribution of the current study.	105
6.11	The waveforms used as the illustration of $b$ -modulation method with their Fourier transforms.	113
6.12	The points available for placing the eigenvalues at, when keeping the exact localisation of $b$ -modulated signal, for different waveforms and scaling parameters.	114
6.13	The signals, generated from the Nyquist waveform via $b$ -modulation with and without additional eigenvalues embedded, for different scaling factors and eigenvalues.	115
6.14	The signals, generated from the flat-top waveform via $b$ -modulation with and without additional eigenvalues embedded, for different scaling factors and eigenvalues.	115
6.15	The effective SNR of the $b(k)$ waveform, arising from the noise in the time domain for different values of the scaling factor.	116
6.16	Schematics of $b$ -modulation transmission system with one additional solitonic mode modulated.	118
6.17	$Q$ -factor dependency on signal power in $b$ -modulated systems with and without embedded (single) eigenvalue, compared with corresponding OFDM transmission.	119

6.18	Schematics of $b$ -modulation transmission system with seven additional solitonic modes modulated. . . . .	119
6.19	$Q$ -factor dependency on signal power in $b$ -modulated systems with and without embedded eigenvalues. . . . .	120
6.20	The illustration of the advantages of PNFT comparably to the (conventional) NFT from an implementation perspective. . . . .	121
6.21	The structure of the NF spectrum of finite-band solution, adjusted to be periodic. . . . .	123
6.22	Normalised power deviation $\sigma_P/P$ as a function of the main spectrum structure for high-genus periodic RHP profiles. . . . .	123
6.23	Isolines of constant bandwidth in zeroth-order approximation in the plane of the number of bands and main spectrum points spacing. . . . .	124
6.24	Two-parametric fit of a deviation of main spectrum point spacing $\delta\Delta\Re\lambda$ as a function of signal period and the number of bands. . . . .	125
6.25	Curves for power manipulation maintaining data rate in space of main spectrum structure parameters. . . . .	126
6.26	The $Q_{\text{EVM}}^2$ -factor of the B2B calculation of phases $\phi_j$ , embedded into the signal, as the function of the cut sampling step $\Delta\Im\lambda$ for the vertical cuts, and time sampling step $\Delta\tau$ . . . . .	127
6.27	The exemplary main spectrum of a genus-1 signal. . . . .	128
6.28	Two examples of an optical signal waveform with the corresponding main spectra images. . . . .	129
6.29	Schematic of the cyclic extension of the periodic signal within the window of the dispersion-caused signal broadening. . . . .	129
6.30	The simulation results of the transmission systems based on genus-1 periodic RHP solution: $Q^2$ -factor as a function of signal power and the exemplary received constellations. . . . .	130
6.31	The simulation results of the transmission systems based on genus-1 periodic RHP solution: $Q^2$ -factor as a function of the link length and the exemplary received constellations. . . . .	131
6.32	The simulation results of the transmission systems based on genus-1 periodic RHP solution: 2D histogram of the received constellation. . . . .	132
6.33	Examples of the generated signal with a fixed 3-cut main spectrum and difference associated phases with the corresponding NF spectrum images. . . . .	133
6.34	Example of the transmitted and received signal, generated from 15-cut RHP main spectrum, carrying 15 8PSK symbols over a link of length $Z = 1080$ km. . . . .	133
6.35	The $Q_{\text{EVM}}^2$ -factor and BER as a function of signal power after $Z = 1040$ km transmission for 15-cuts RHP-based transmission system. . . . .	134
6.36	The received phase constellation after 1000 km transmission of 8PSK auxiliary spectrum for 15-cuts RHP-based transmission system. . . . .	134
6.37	The $Q_{\text{EVM}}^2$ -factor as a function of transmission distance at optimum signal power $P = -17$ dBm. . . . .	135
B.1	NLS solution via NFT - I . . . . .	151
B.2	NLS solution via NFT - II . . . . .	152
B.3	NFT-based transmission - I . . . . .	152
B.4	NFT-based transmission - II . . . . .	152
B.5	NFT-based transmission - III . . . . .	153
B.6	NFT-based soliton transmission . . . . .	153
B.7	NF spectrum for vanishing boundary conditions . . . . .	154
B.8	PNFT vs (vanishing) NFT on Tx and Rx - I . . . . .	154

B.9 PNFT vs (vanishing) NFT on Tx and Rx - II . . . . .	155
B.10 PNFT vs (vanishing) NFT on Tx - I . . . . .	155
B.11 PNFT vs (vanishing) NFT on Tx - II . . . . .	155
B.12 PNFT vs (vanishing) NFT full comparison . . . . .	156
B.13 Details of NF domain - time domain mapping . . . . .	157
B.14 Transfer matrices approach to ZSP solution . . . . .	157

# PhD publications

In this thesis are presented results, published or submitted to publication in a sequence of scientific papers and topical conference contributions. These publications are listed below:

- J1 Morteza Kamalian, **Anastasiia Vasylichenkova**, Dmitry Shepelsky, Jaroslaw E. Prilepsky, and Sergei K. Turitsyn, Signal Modulation and Processing in Nonlinear Fibre Channels by Employing the Riemann–Hilbert Problem in *Journal of Lightwave Technology*, vol. 36, no. 24, pp. 5714-5727 (2018).
- J2 **Anastasiia Vasylichenkova**, Jaroslaw E. Prilepsky, Dmitry Shepelsky, and Amit Chattopadhyay, Direct nonlinear Fourier transform algorithms for the computation of solitonic spectra in focusing nonlinear Schrödinger equation // *Communications in Nonlinear Science and Numerical Simulation* 68C, pp. 347-371 (2019).
- J3 **Anastasiia Vasylichenkova**, Jaroslaw E. Prilepsky, and Sergei K. Turitsyn, Contour integrals for numerical computation of discrete eigenvalues in the Zakharov–Shabat problem // *Opt. Lett.* 43, 3690-3693 (2018).
- J4 Dmitry Shepelsky, **Anastasiia Vasylichenkova**, Jaroslaw E. Prilepsky, and Iryna Karpenko, Nonlinear Fourier spectrum characterization of time-limited signals // in *IEEE Transactions on Communications* 68, 5 (2020).
- J5 Maryna Pankratova, **Anastasiia Vasylichenkova**, Stanislav A. Derevyanko, Nikolay B. Chichkov, and Jaroslaw E. Prilepsky, Signal-noise interaction in optical fiber communication systems employing nonlinear frequency division multiplexing // in *Phys. Rev. Applied* 13, 054021 (2020).
- SJ1 Morteza Kamalian, **Anastasiia Vasylichenkova**, Dmitry Shepelsky, Jaroslaw E. Prilepsky, and Sergei K. Turitsyn, Full-spectrum periodic nonlinear Fourier transform optical communication through solving the Riemann-Hilbert problem // accepted to *Journal of Lightwave Technology* (2020).
- SJ2 Oleksandr Kotlyar, Maryna Pankratova, Morteza Kamalian, **Anastasiia Vasylichenkova**, Jaroslaw E. Prilepsky, and Sergei K. Turitsyn, “Combining nonlinear Fourier transform and neural network-based processing in optical communications”, accepted to *Optics Letters* (2020).
- C1 **Anastasiia Vasylichenkova**, Jaroslaw E. Prilepsky, Nikolai B. Chichkov, Sergei K. Turitsyn, Multieigenvalue communication paired with b-modulation // in *European Conference of Optical Communication* (Dublin, Ireland, Sept 2019).
- C2 Morteza Kamalian-Kopae, **Anastasiia Vasylichenkova**, Dmitry Shepelsky, Jaroslaw E. Prilepsky, Sergei K. Turitsyn, Multisymbol Periodic Nonlinear Fourier Transform Communication // in *European Conference of Optical Communication* (Dublin, Ireland, Sept 2019).

- C3 **Anastasiia Vasylychenkova**, Morteza Kamalian, Dmitry Shepelsky, Jaroslaw E. Prilepsky, Sergei K. Turitsyn, Periodic Nonlinear Fourier transform for the optical fibre nonlinearity mitigation via solution of Riemann-Hilbert Problem // Proceedings of 3rd IMA Conference ‘Nonlinearity and Coherent Structures’, Newcastle, UK, 2019.
- C4 **Anastasiia Vasylychenkova**, Jaroslaw E. Prilepsky, Nikolai B. Chichkov, and Sergei K. Turitsyn, Combining the Discrete NFT Spectrum with b-modulation for High-Efficiency Optical Transmission // in 2019 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC), Munich, Germany, Paper CI-2.5.
- C5 Morteza Kamalian Kopae, **Anastasiia Vasylychenkova**, Jaroslaw E. Prilepsky, Dmitry Shepelsky, Sergei K. Turitsyn. Communication System Based on Periodic Nonlinear Fourier Transform with Exact Inverse Transformation // in European Conference of Optical Communication (Roma, Italy, Sept 2018), paper Tu3A.2.
- C6 Morteza Kamalian Kopae, Dmitry Shepelsky, **Anastasiia Vasylychenkova**, Jaroslaw E. Prilepsky, Sergei K. Turitsyn. Communication System Using Periodic Nonlinear Fourier Transform Based on Riemann-Hilbert Problem // in European Conference of Optical Communication (Roma, Italy, Sept 2018), paper Tu3A.4.
- C7 Morteza Kamalian, Jaroslaw E. Prilepsky, **Anastasiia Vasylychenkova**, Dmitry Shepelsky, Sergei K. Turitsyn. Methods of nonlinear Fourier-based optical transmission with periodically-extended signals // at 2018 ICSEE International Conference on the Science of Electrical Engineering, Israel, Dec 2018.
- C8 **Anastasiia Vasylychenkova**, Jaroslaw E. Prilepsky, Amit Chattopadhyay, Nonlinear Fourier Transformation: discrete spectrum methods analysis // Proceedings of 2nd IMA Conference ‘Nonlinearity and Coherent Structures’, Norwich, UK, 2017.
- C9 **Anastasiia Vasylychenkova**, Comparison of Nonlinear Fourier Transform methods performance and accuracy // Proceedings of ‘13th Developing in Optics and Communications Conference’, Riga, Latvia, 2017.

During the PhD study, I have been involved in various studies within the NFT research group at AiPT, which also have been published. I am not including the results from those to my thesis, however, I cite them through the text and I would like to mention my additional contribution to the field of NFT-based communication, besides the specific topic of my thesis. These publications are listed below:

- CC1 **Anastasiia Vasylychenkova**, Maryna Pankratova, Jaroslaw E. Prilepsky, Nikolai Chichkov, and Sergei K. Turitsyn, Signal-dependent noise for b-modulation NFT-based transmission // in 2019 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC), Munich, Germany, Paper CI-1.2.
- CC2 Morteza Kamalian Kopae, **Anastasiia Vasylychenkova**, Oleksandr Kotlyar, Maryna Pankratova, Jaroslaw E. Prilepsky, and Sergei K. Turitsyn, Artificial Neural Network-Based Equaliser in the Nonlinear Fourier Domain for Fibre-Optic Communication Applications // in 2019 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC), Munich, Germany, Paper CI-1.4.

- CC3 Oleksandr Kotlyar, Maryna Pankratova, Morteza Kamalian, **Anastasiia Vasylichenkova**, Jaroslaw E. Prylepskiy, Sergei K. Turitsyn, Unsupervised and supervised machine learning for performance improvement of NFT optical transmission // in 2018 IEEE British and Irish Conference on Optics and Photonics (BICOP), London, United Kingdom, 2018.
- CC4 Maryna Pankratova, **Anastasiia Vasylichenkova**, Jaroslaw E. Prylepskiy, Stanislav A. Derevyanko, Study of Noise-Induced Signal Corruption for Nonlinear Fourier-Based Optical Transmission // in 2018 IEEE British and Irish Conference on Optics and Photonics (BICOP), London, United Kingdom, 2018.
- CC5 Maryna Pankratova, **Anastasiia Vasylichenkova**, Jaroslaw E. Prilepsky, and Stanislav A. Derevyanko, Properties of the effective noise in the nonlinear Fourier transform-based transmission // in proceedings of Frontiers in Optics / Laser Science, OSA Technical Digest (Optical Society of America, 2018), paper JW3A.83.
- CC6 Maryna Pankratova, **Anastasiia Vasylichenkova**, Jaroslaw E. Prilepsky, Noise-induced signal corruption in nonlinear Fourier-based optical transmission system in the presence of discrete eigenvalues // in The International Conference on Optical Communication Systems (Prague, Czech Rep, July 2019).
- CC7 Stanislav Derevyanko, Dmitry Shepelsky, Maryna Pankratova, **Anastasiia Vasylichenkova**, Nikolai Chichkov, and Jaroslaw Prilepsky, Analytical model of nonlinear noise in the b-modulated optical transmission systems // Proc. 2020 Conference on Lasers and Electro-Optics, paper SF2L.5 (CLEO, San Jose, US, May 2020).

# Notations and list of acronyms

The general idea I follow throughout the text: normal text letters mean scalar quantities, lower-case bold letters - vectors, columns, etc., upper-case bold letter - matrices (except Pauli matrices  $\sigma_{1-3}$ , eq. (1)), hatted bold letters - operators. The normalised quantities are usually designated as a lower-case version of the letter for a not-normalised (in physical units) counterpart. A monospaced font is used for programming commands, variables etc., in the main text, appendix and code listings.

The matrices entries and columns are involved in the expressions apart from the matrices: the individual cells are marked with double-indexing in brackets ( $\mathbf{A}_{(12)}$  for the right-top entry) and the columns are designated with single-indexing in brackets ( $\mathbf{A}_{(1)}$  for the left column).

The sets are designated with `mathcal` command, excluding  $\mathcal{N}$  used for the number of gaps for Riemann-Hilbert problem input (in chapters 3 and 6). The common sets are designated with `mathbb` command:  $\mathbb{R}$  for real numbers,  $\mathbb{Z}$  for integers,  $\mathbb{C}$  for complex numbers. In addition, the (closed) half-planes are:  $\mathbb{C}_{\pm} = \{k \in \mathbb{C} : \pm \Im k > 0\}$  and  $\overline{\mathbb{C}}_{\pm} = \{k \in \mathbb{C} : \pm \Im k \geq 0\}$ .

In the set of complex numbers,  $i$  designates imaginary unit,  $\Re$  and  $\Im$  stay for real and imaginary part, correspondingly. Asterics  $a^*$  means the complex conjugation:  $a^* = \Re a - i \Im a$ .

The definitions of important entries used in the text are the following. The Pauli matrices are defined as:

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1)$$

The Dirac delta function  $\delta(x)$  is a generalised function, limiting the rectangular function to the infinite amplitude and zero support, used as localising weight and normalised to the unity, as given:

$$\int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx = f(x_0), \quad \int_{-\infty}^{\infty} \delta(x) dx = 1. \quad (2)$$

I use the (normalised) definition of sinc function, as given:

$$\text{sinc}(x) := \frac{\sin(x)}{x}. \quad (3)$$

The diagonal matrix is uniquely determined by the (main) diagonal entries, with all remaining entries being zero:

$$\text{diag}(a, b, c, d) := \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{pmatrix}. \quad (4)$$

The Toeplitz matrix is a diagonal-constant matrix, uniquely determined by the entries of first column (and row):

$$\text{Toeplitz}(a, b, c, d; a, e, f, g) := \begin{pmatrix} a & e & f & g \\ b & a & e & f \\ c & b & a & e \\ d & c & b & a \end{pmatrix}. \quad (5)$$

In the matrix (and vector) operations, superscript  $\mathbf{A}^T$  means transposition, whilst superscript  $\mathbf{A}^+$  stands for Hermitian conjugation.

I use standard deviation as useful matrix for stochastic quantities deviation, for quantity  $Y$ , having mean value  $\langle Y \rangle$ , the standard deviation is:

$$\sigma_Y^2 = \langle (Y - \langle Y \rangle)^2 \rangle, \quad (6)$$

and normalised deviation is  $\sigma_Y / \langle Y \rangle$ .

Single vertical bracketing  $|a|$  means absolute value, double vertical bracketing for vectors  $\|\mathbf{v}\|$  means Euclidean norm given as:

$$\|\mathbf{v}\|^2 = \sum_j v_{(j)}^2, \quad (7)$$

square bracketing  $[a]$  means integer part of the number (rounding ignoring part after decimal point), angle bracketing  $\langle \rangle$  means averaging over the relevant selection set.

In notations of physics effects and objects, I follow conventional practices, meaning  $t$  as time,  $x, y, z$  as coordinates,  $f$  and  $\omega$  are frequency and circular frequency (for waves),  $n$  as refractive index,  $c$  as the speed of light, etc. I do not provide the full listing of notation, I hope that the meaning will be clear from the context.

Following the widely used acronyms, l/r.h.p. means left/right-hand side part of the equation. The full list of used acronyms with the page of the first appearance is provided below:

<b>AKNS</b> Ablowitz, Kaup, Newell, Segur.....	37
<b>AL</b> Ablowitz-Ladik.....	67
<b>ASE</b> amplifier spontaneous emission.....	25
<b>ASIC</b> application-specific integrated circuit.....	30
<b>AWGN</b> additive white Gaussian noise.....	33
<b>BER</b> bit error rate.....	31
<b>BO</b> Bofetta-Osborne.....	66
<b>CN</b> Crank-Nikolson.....	67

<b>DBP</b> digital backpropagation.....	30
<b>DL</b> Delves and Lyness.....	83
<b>DT</b> Darboux Transfrom .....	95
<b>EDFA</b> Erbium-dopped fibre ampliffier.....	24
<b>EVM</b> error-vector magnitude.....	31
<b>FC</b> Fourier collocation.....	78
<b>FT</b> Fourier Transform.....	27
<b>GLM</b> Gelfand-Levitan-Marchenko equation.....	50
<b>GVD</b> group-velocity dispersion.....	22
<b>IFT</b> Inverse Fourier Transform .....	27
<b>INFT</b> inverse Nonlinear Fourier Transfrom.....	35
<b>IST</b> Inverse Scattering Transform.....	35
<b>KdV</b> Korteweg-de Vries .....	35
<b>ML</b> machine learning.....	30
<b>NF</b> Nolinear Fourier.....	35
<b>NFT</b> Nonlinear Fourier transfrom .....	19
<b>NLS</b> Nonlinear Schrödinger equation.....	25
<b>NR</b> Newton-Raphson.....	75
<b>OFDM</b> orthogonal frequency-division multiplexing.....	99
<b>OPC</b> optical phase conjugation.....	29

**PAPR** peak-to-average power ratio ..... 82

**PNFT** periodic Nonlinear Fourier transform ..... 97

**RHP** Riemann-Hilbert problem ..... 44

**ROI** region of interest ..... 78

**Rx** receiver ..... 30

**SE** spectral efficiency ..... 29

**SMF** single-mode fibre ..... 22

**SNR** signal-to-noise ratio ..... 29

**Tx** transmitter ..... 31

**ZSP** Zakharov-Shabat problem ..... 37

# Chapter 1

## Introduction

### 1.1 Motivation

The improvement of optical communication technologies has been growing in the alignment with the data traffic demand, caused by continuous development of other economic sectors. The modern straightforward and linear technologies cannot properly fulfil the anticipated needs in the achievable data rates and propagation distance.

Depending on the target reach, the limiting factors and their contribution may vary. The fibre nonlinearity is considered as an important bottleneck for long-haul transmissions, because of operating in high-power and long-distance regime.

The Nonlinear Fourier transform (NFT) is studied in this thesis as a methodology for effective signal processing, which effectively linearises the optical channel. The processing is aligned on both transceiver and receiver. In this concept, the signal is (de)composed to the natural mode, which propagates independently in the optical fibre. This method involves the advances of mathematical physics and needs to be adapted to the application for communication. This adaptation includes the study of the sampling requirements, development of effective and fast numerical methods. Additionally, the research is required for the development of new transmission formats, making and motivating the decisions about the manipulation with signal parameters and degrees of freedom for information coding.

### 1.2 Thesis outline

The remaining part of the thesis consists of 6 chapters, from chapter 2 to chapter 7.

Chapter 2 is dedicated to the survey and review of the typical practices, most important and most recent achievements in the subject area. It includes the review of the physical effects, occurring in the optical fibre, the most relevant for the convenient optical communications application, and describes the comprehensive channel model with relevant components. The model is discussed from the analytical and numerical study point of view. Additionally, it lists main nonlinearity mitigation methods and details of simulations of the transmission systems.

Chapter 3 contains the complete description of the NFT (IST) approach, from the mathematical point of view. Two possible boundary conditions cases (vanishing and periodic) are equivalently addressed there. This chapter includes the details of the correspondence between operators, definitions of objects and their properties, scattering data structures and their evolution in time. The mathematical procedures for both direct and inverse NFT, for both vanishing and periodic boundary conditions are presented there. In the case of several possible approaches (to the same transform), their advantages and applicability are compared and discussed.

Chapter 4 presents the analysis of the performance of direct NFT methods, basing of analytical and pre-constructed examining cases. It is focused on the numerical accuracy and computational time of the existing methods. Also, it provides the derivation and testing of novel improved numerical methods. Cases for both continuous and discrete scattering data are presented.

In chapter 5, for the first time, I fully present my results in the developing of the novel numerical method for discrete eigenvalues evaluation. These results have been published only fragmentarily before. Among the advantages of this method is simultaneous detection of all embedded eigenvalues and the predictable numerical complexity. The testing and comparison with the existing discrete eigenvalue methods are provided, both for analytical and pre-constructed signals. Aiming the practical implementation, for these methods the noise tolerance is studied.

Chapter 6 contains the results in the developing of NFT-based transmission systems. The advantages and limitation of both boundary condition cases are studied and discussed.

Finally, in chapter 7, I conclude the overall results of the research work and discuss the perspectives of the NFT method for nonlinearity mitigation.

### 1.3 Contribution

In this section, I want to properly acknowledge the collaboration and contribution during this research work.

I have two main collaborators, who worked with me on the results presented in this thesis, related to the periodic boundary condition case. Our professional expertise is complementary, which allowed the effective and successful teamwork. The first collaborator is Dr Morteza Kamalian, who has communication background, so he is more responsible for the numerical simulation of the transmission systems. Another collaborator is Prof Dmitry Shepelsky, he is more responsible for the analytical derivations, Riemann-Hilbert problems equivalent deformations and for justification of the general Riemann-Hilbert problem concept. I contributed to all stages of the periodic systems developments, with their guidance and support.

In vanishing boundary condition research, I have developed the numerical realisations and simulations myself, under my supervisors' guidance.

Additionally, I have to mention the overall supervision and management, done by my main supervisor, Prof Sergei Turitsyn, and associate supervisor, Dr Yaroslav Prilepskiy. They contributed mainly during the papers preparation stages and by general guiding and advising.

### 1.4 Tools

The numerical realisations of algorithms and numerical simulations of the transmission systems I mainly implemented in `MATLAB`. During the study, I used different versions of `MATLAB` from 2016b to 2019a, available under the Aston University licence. Some parts of simulations, related to the Riemann-Hilbert problem, were implemented in Wolfram Mathematica to avoid data transfer between different software. I used Wolfram Mathematica versions 10-11, available under the Aston University licence.

In the last part of my PhD research, I used fast and parallelised implementation of the split-step numerical algorithm, developed by Dr Nikolai Chichkov, one of my co-authors. It is roughly related to the period of publications [C1, C4].

For the fast implementation of the NFT numerical method, I used the open-access code, runnable from `MATLAB`, developed by the research group of Wahls [139]. For the

numerical solutions of the Riemann-Hilbert problem, I used the open-access numerical package in Wolfram Mathematica by Olver [107].

This thesis does not contain any experimental results, so I do not provide any information about experimental tools and equipment here.

# Chapter 2

## Current state of art

### 2.1 Physical effects in the optical fibre

#### 2.1.1 General remarks

The very idea of the optical fibre lies in the locking light inside the directed thin glass tube due to the effect of total internal reflection. This is achieved by creating a two-layer tube, having an internal part (core) with a higher refractive index than an external part (cladding). Depending on the cross-section area of the core, there is some freedom in the beam's directions between reflections. For sufficiently thin fibres, all beams are concentrated longitudinally to the fibre axis. Such fibres, with the limited variety of light directions, are called *single-mode* fibres, meaning as a mode the possible light path in the fibre's core.

In this section, I provide the listing of the physical effect occurring in the optical fibre. I do not mention all possible effects, limiting to the minimally required, so I mention only these phenomena which are relevant to (i) fibre-optics in application to optical communications, and (ii) typical long-distance application, in particular, the standard single-mode fibre (SMF). The application of optical fibres goes much beyond these two areas, including technologies for designing several-modes and several-core fibres, grading-index fibres, etc. [4]. The optical fibres are also involved for the dispersion-management, fibre lasers and amplifiers, however, in these areas, the focus is shifted to modifying the fibre structure or content. For other, less relevant effects, the reader can refer to one of fibre optics classic book by Agrawal [4].

Note that I intend to provide the measurement units for the parameters in SI.

#### 2.1.2 Group-velocity dispersion

One of the main effects, influencing the light propagation in the optical fibre arises from the refractive index  $n$  dependency on the wavelength (or, equivalently, frequency). Consequently, the same dependency arises for the wavenumber  $\beta$ . In the assumption of a thin spectrum of the beam, almost monochromatic wave at wavelength  $\lambda_0$  (so at circular frequency  $\omega_0$ ), this dependency can be decomposed to the Taylor series as:

$$\beta(\omega) = \frac{n(\omega)\omega}{c} = \sum_{m=0}^{\infty} \frac{\beta_m}{m!} (\omega - \omega_0)^m = \beta_0 + \beta_1(\omega - \omega_0) + \frac{1}{2}\beta_2(\omega - \omega_0)^2 + \dots \quad (2.1)$$

The first-order term can be eliminated by considering the co-moving (with the pulse) reference frame, which travels with the pulse group velocity  $v_g$  (velocity with which the overall envelope shape of the wave's amplitudes propagates through space). The second-order term represents the dependency of the group velocity on the wavelength, so is called group-velocity dispersion (GVD).

GVD can be alternatively characterized by the dispersion parameter:

$$D = \left. \frac{d}{d\lambda} \frac{1}{v_g} \right|_{\lambda=\lambda_0} = -\frac{2\pi c}{\lambda_0^2} \beta_2, \quad (2.2)$$

measured in  $\left[\frac{\text{s}}{\text{m}^2}\right]$ .

Practically, at least for smooth single-lobe profiles, the presence of GVD leads to the signal broadening: its effective time-support expands during the propagation. It can be naively explained in the way that the shorter-wavelength components of the pulse delay in comparison with the longer ones (for silica,  $D > 0$ ). For the known signal bandwidth  $B$  (in Hz), after the propagation over the distance  $Z$ , the overall signal time-support expansion can be estimated as:

$$\Delta t = 2\pi |\beta_2| Z B. \quad (2.3)$$

This gap estimation is known also as the channel *dispersion-induced memory* and used to determine the sufficient separation between individual signal bursts.

The expansion above allows accounting higher-order dispersions, however, they start being relevant for shorter pulses, more applicable for laser physics than to communications. The third-order dispersion is characterised by the *dispersion slope*  $S$ , measured in  $\left[\frac{\text{s}}{\text{m}^3}\right]$ .

### 2.1.3 Nonlinear effects

The main relevant effect is the so-called *Kerr nonlinearity*, expressing as the dependency of the core refractive index  $n$  on the light intensity, and in the assumption of low nonlinearity, it can be represented as an expansion:

$$n = n_0 + n_2 |E|^2. \quad (2.4)$$

Due to the symmetry of  $\text{SiO}_2$ , the first-order term is not involved in this expansion. In silica fibres, the nonlinear contribution to polarisation vector  $\mathbf{P}$  (by molecular vibration) is viewed as a small perturbation and in the approximation of the electric-dipole interaction, the medium response is local both in position and time.

For the following derivations of the wave amplitude (envelope) evolution, it is convenient to characterise the nonlinearity via the nonlinearity coefficient as:

$$\gamma = \frac{n_2 \omega_0}{c A_{\text{eff}}}, \quad (2.5)$$

where  $A_{\text{eff}}$  is an effective core area, determined by the spatial distribution of the light over the fibre cross-section. The measurement unit is  $[1/\text{W}/\text{m}]$ .

The Kerr nonlinearity is responsible for self-phase and cross-phase modulations, when the nonlinearity-caused phase deviation is specifically addressed. However, in this research, the Kerr effect is considered as a part of the signal propagation model, not separating its effects on the signal amplitude and phase.

### 2.1.4 Loss and amplification

The light-medium interaction inevitably leads to an energy transfer from the wave to the material. In this way, the medium is heating and the pulse is losing its power. The energy (power) loss during the pulse propagation in the fibre is characterised by the loss coefficient  $\alpha$ , in the simplified approach defined as:

$$\frac{\partial P}{\partial z} = -\alpha P, \quad (2.6)$$

so the power  $P$  is decaying exponentially with factor  $\alpha$ , when the electric field  $E$  is decaying with factor  $\alpha/2$ . The natural unit for loss coefficient is  $[1/\text{m}]$ , however, it is widely

expressed in [dB/m], to deal with the signal power, measured in dBm, in a more convenient way.

The pulse bandwidth is assumed to be far from any material resonances. For the transparency window of silica glass, the dependency of the attenuation coefficient of the wavelength is provided in fig. 2.1. As it is labelled there, the main contribution to the power loss in the optical fibre is caused by Rayleigh scattering (left part, decaying as  $\lambda^{-4}$ ) and material absorption (right rising part). Also, because of the specific condition of fibre fabrication, some amount of water molecules is left in the fibre core. This spectrum line in the left part of the loss coefficient dependency is known as OH-peak.



Figure 2.1: The attenuation coefficient of silica fibre near the transparency point as a function of the incident wave frequency (cited from [122]).

In modern communications, the band around the extremum wavelength  $\lambda_0 = 1550$  nm is used, exploiting approximately 5 THz bandwidth. This frequency range is known as a conventional band or C-band. Within this range, the attenuation is considered to be frequency-independent, at the value  $\alpha = 0.2$  dB/km.

Two common approaches are used for loss compensation in the optical fibre. First, I have to note that the rates of currently employed systems require optical solutions in the signal amplification: converting to the electrical domain and back would be more expensive and would limit the system performance. The game-changing discovery of Erbium-doped fibre amplifier (EDFA) allowed the sensational expansion of optical communication capacity in the 1990s [15, 33]. Adding and exciting (with laser pump) the  $\text{Er}^{3+}$  ions in silica enables to effectively reproduce the incident photons, operating exactly in the range of silica transparency window. In this approach, the amplifiers are located discretely, on usually 80 km spans, so, this kind of amplifiers boosts the signal locally, compensating the exact amount of energy, lost during one span propagation.

Another common way to maintain the signal power during its propagation through the fibre is to use (distributed) Raman amplifiers. There are several approaches to Raman amplifier design, using forward or backward laser pump to excite the gain medium [15]. The Raman amplification provides almost constant gain profile along the span [8, 9, 15, 57, 77]. In this thesis, I consider that Raman amplification fully compensates loss in the fibre on the length of the span. In practice, Raman amplifiers provide not-constant profile, consisting of convex and concave humps, their order depends in the direction of the amplification pump (either forward or backward). Because of the variety of designs, there is no preferable configuration, and therefore, no standard for a gain-loss profile in the presence of distributed Raman amplification. This distribution could be taken into account in the simulations, however, with marginal effect on the transmission performance,

because of many more impactful factors.

The presence of the amplifiers inevitably causes another and the most important noise source for long-distance transmission, the so-called amplifier spontaneous emission (ASE) noise. When the pump laser pulse passes the gain medium, it can cause uncontrollable and random electron transitions. For more details to analytical models of this noise see subsection 2.6.4.

## 2.2 Optical fibre modelling with Nonlinear Schrodinger equation

The concept of NFT-based optical communications, in any case, must be related with involvement of so-called *integrable* (or almost integrable) equation as an effective model of the signal propagation. From the physics point of view, this model must include effects of the interaction between the electric field and the optical fibre media, taking into account the effects from the subsections above.

Starting with the general wave equation for the electric field, originating from Maxwell's equations for vector electric field  $\mathbf{E}$  and vector polarisation  $\mathbf{P}$ :

$$\nabla \times \nabla \times \mathbf{E} = -\frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2}. \quad (2.7)$$

The subsequent derivation accounts several assumptions, related to the properties of the fibre. They include the involvement of axial symmetry, which allows reducing the vector equation to the scalar form, considering only the coordinate along the fibre  $z$ . Accounting the symmetry of  $\text{SiO}_2$  molecule simplifies the material (electromagnetic) permittivity from tensor form to scalar. The nonlinear material response is considered to be perturbative comparably to the linear response. The propagating light is assumed to be quasi-monochromatic, so instead of describing the evolution of the electric field, the problem is reduced to the amplitude evolution. The main (central) wavelength is chosen according to the silica transparency window, see fig. 2.1, to be 1550 nm.

These assumptions and simplifications allow writing the equation for the slow-varying envelope of the electric field  $A(t, z)$ , passing through the optical fibre as:

$$i \frac{\partial A}{\partial z} = -i \frac{\alpha}{2} A + \frac{\beta_2}{2} \frac{\partial^2 A}{\partial t^2} - \gamma |A|^2 A. \quad (2.8)$$

Note here that  $t$  designates the time in the co-moving reference frame, so, can be negative, whilst  $z$  is the distance along the fibre and only positive. This equation is commonly referred as Nonlinear Schrödinger equation (NLS), however in mathematics NLS usually depicts the evolutionary equation with dispersion (subsection 2.1.2) and third-order nonlinearity (subsection 2.1.3) terms only:

$$i \frac{\partial A}{\partial z} = \frac{\beta_2}{2} \frac{\partial^2 A}{\partial t^2} - \gamma |A|^2 A. \quad (2.9)$$

Note that in the following work I assume that the fibre has anomalous dispersion ( $\beta_2 < 0$ ), what leads to the *focusing* NLS. During the propagation, GVD and Kerr nonlinearity compete, decreasing the beam effective cross-section.

Introduction of additional terms allows the generalisation of NLS taking into account other effects in the optical fibre, i.e. higher-order dispersion or nonlinearity, including stimulated scattering [4]. Besides, it ignores the existence of two polarisations of light, meaning that the polarization is maintained during the propagation. In this thesis, I do not move beyond this most common form of the channel model, to use the properties of its integrability for the sake of optical communications.

Note that the two-polarisation generalisation of NLS exists and is known as Manakov equation [4, 98], where this approach is applicable. For two-polarisation case, the power transfer between polarisation states is taken into account. The study of this case and applicability of NFT approach to it is put of the scope of this thesis, but some study for periodic NFT can be found in [74, 118, 119].

### 2.2.1 Path averaging

The NLS in the form eq. (2.8) is not integrable, so we cannot apply NFT to it for solving or for coding purposes. To be able to use NFT for NLS with losses, one can use path averaging techniques, see [64, 85]. Redefining variables as  $A := \exp(-\alpha/2 \cdot z)Q$ , one can get the following:

$$i \frac{\partial Q}{\partial z} = \frac{\beta_2}{2} \frac{\partial^2 Q}{\partial t^2} - \gamma \exp(-\alpha z) |Q|^2 Q. \quad (2.10)$$

Principally, this equation represents the propagation of the signal  $Q(z, t)$  in the medium with effectively changing nonlinearity coefficient as  $\gamma \exp(-\alpha z)$ . Assuming the case, when the sap length is much less than all characteristic lengths of the fibre, one can introduce an effective nonlinearity coefficient along the fibre (or span) of the length  $L_{\text{sp}}$  as:

$$\gamma_{\text{eff}} = \frac{1}{L_{\text{sp}}} \int_0^{L_{\text{sp}}} \exp(-\alpha z) \gamma \, dz = \gamma \frac{1 - \exp(-\alpha L_{\text{sp}})}{\alpha L_{\text{sp}}}. \quad (2.11)$$

The same expression is often written in terms of the amplifier gain as [85]  $\gamma_{\text{eff}} = \gamma(G - 1)/(G \log(G))$ , where  $G = \exp(\alpha L_{\text{sp}})$ .

Replacing  $z$ -dependent nonlinearity with the effective constant one in eq. (2.10), one can get

$$i \frac{\partial Q}{\partial z} = \frac{\beta_2}{2} \frac{\partial^2 Q}{\partial t^2} - \gamma_{\text{eff}} |Q|^2 Q, \quad (2.12)$$

what is literally purely integrable NLS in the form (2.9).

### 2.2.2 Normalisation approaches

There are two widely used ways to normalise eq. (2.9), which lead to different coefficients in front of the principal terms. This also influences some details of the corresponding inverse scattering problem.

All normalisations are uniquely determined by a single normalisation parameter, in traditions of our research group I start with normalisation time  $T_0$ , however, one can choose to start either from distance scale or power scale as well:

- (i) Having  $T_0$ , one can define  $L_0 = T_0^2/|\beta_2|$  and  $P_0 = 1/\gamma/L_0 = |\beta_2|/\gamma/T_0^2$ . Then normalise the temporal variable as  $t = T_0\tau$ , the space variable as  $z = L_0\zeta$  and the signal function as  $Q = \sqrt{P_0}q$ . It leads to the equation:

$$i \frac{\partial q}{\partial \zeta} = \frac{1}{2} \frac{\partial^2 q}{\partial \tau^2} + |q|^2 q, \quad (2.13)$$

or in presence of the loss term:

$$i \frac{\partial q}{\partial \zeta} = -i \frac{\alpha_0}{2} q + \frac{1}{2} \frac{\partial^2 q}{\partial \tau^2} + |q|^2 q \quad (2.14)$$

with  $\alpha_0 = \alpha L_0 = \alpha T_0^2/|\beta_2|$  acting as a normalised attenuation parameter.

- (ii) Having the same normalisation time  $T_0$ , one can define  $L_0 = 2T_0^2/|\beta_2|$  and  $P_0 = |\beta_2|/\gamma/T_0^2$  and normalise the variables as  $t = T_0\tau$ ,  $z = L_0\zeta$  and  $Q = \sqrt{P_0}q$ . The only difference here lies in different normalisation of space variable. The normalised NLS in this case has a form:

$$\frac{\partial q}{\partial \zeta} = i \frac{\partial^2 q}{\partial \tau^2} + 2i|q|^2q, \quad (2.15)$$

or, taking loss into account:

$$\frac{\partial q}{\partial \zeta} = -\frac{\alpha_0}{2}q + i \frac{\partial^2 q}{\partial \tau^2} + 2i|q|^2q, \quad (2.16)$$

with two times larger normalised attenuation in that case  $\alpha_0 = \alpha L_0 = 2\alpha T_0^2/|\beta_2|$ .

Here I note that during this research I tried to stick to the normalisation (ii) provided. However, some of the results are invariant to the normalisation, as, e.g. numerical methods accuracy study from chapter 4.

## 2.3 Numerical solution of initial value problem for NLS

The signal propagation in the optical fibre is principally modelled by means of NLS, including the amplification scheme, either lumped EDFA, or distribute Raman amplification. For purposes of numerical studies, the signal propagation in the fibre is simulated numerically by split-step Fourier scheme [4].

To present the input signal on a finite support one uses an array of equidistant samples  $\mathbf{f}$ . The propagation distance is also discretised onto equal steps. To make it more universal, I write the algorithm for NLS for unknown function  $f(l, \vartheta)$ , where  $l$  may be understood as a distance variable, and  $\vartheta$  as a temporal one, with arbitrary coefficients  $C_{\text{loss}}$ ,  $C_{\text{disp}}$ ,  $C_{\text{nonl}}$ , derived elsewhere:

$$\frac{\partial f(l, \vartheta)}{\partial l} = \left[ -C_{\text{loss}}f(l, \vartheta) + iC_{\text{disp}} \frac{\partial^2 f}{\partial \vartheta^2} \right] + \left[ iC_{\text{nonl}}|f|^2f \right] = \hat{\mathbf{D}}f(l, \vartheta) + \hat{\mathbf{N}}f(l, \vartheta). \quad (2.17)$$

The idea of the split-step approach lies in splitting the r.h.p. of eq. (2.17) onto the linear  $\hat{\mathbf{D}}$  and nonlinear  $\hat{\mathbf{N}}$  operators. The evolution over the single discretisation step is performed in three stages: first, half-step linear evolution is performed, then nonlinear evolution, finally, another half-step linear evolution. Elementary evolution is done using matrix exponents: linear half-steps in the (linear) frequency domain and nonlinear step in the time domain.

For the transition between time and frequency domains one can use Fourier Transform (FT), MATLAB command `fft`, and its inverse counterpart Inverse Fourier Transform (IFT), MATLAB command `ifft`. The application of fast (Cooley-Tukey) algorithm for FT allows reaching high speed of simulation [27]. However, because of application of these algorithms, in split-step simulation we unconsciously assume that NLS boundary conditions are periodic, what should be considered properly.

The scheme the single iteration of the split-step algorithm can be written as (here  $\Delta l$  is a step along with the spacial variable  $l$ ):

$$\begin{aligned} & \left[ \text{signal } \mathbf{f} \xrightarrow{FT} \right] \text{spectrum } \hat{\mathbf{f}} \rightarrow \hat{\mathbf{f}} \times \exp \left( \hat{\mathbf{D}} \frac{\Delta l}{2} \right) \xrightarrow{FT} \mathbf{f} \rightarrow \\ & \rightarrow \mathbf{f} \times \exp \left( \hat{\mathbf{N}} \Delta l \right) \xrightarrow{FT} \hat{\mathbf{f}} \rightarrow \hat{\mathbf{f}} \times \exp \left( \hat{\mathbf{D}} \frac{\Delta l}{2} \right) \left[ \xrightarrow{IFT} \mathbf{f} \right] \end{aligned}$$

Table 2.1: Listing of the split-step coefficients for different forms of optical fibre models based on NLS.

Equation/coefficient	$C_{\text{loss}}$	$C_{\text{disp}}$	$C_{\text{nonl}}$
not-normalised NLS w/o loss (2.9)	0	$\beta_2/2$	$\gamma$
not-normalised NLS w/ loss (2.8)	$\alpha/2$	$\beta_2/2$	$\gamma$
normalised NLS (i) w/o loss (2.13)	0	1/2	1
normalised NLS (i) w/ loss (2.14)	$\alpha_0/2$	1/2	1
normalised NLS (ii) w/o loss (2.15)	0	1	2
normalised NLS (ii) w/ loss (2.16)	$\alpha_0/2$	1	2

Inside the loop, it is practical to remove IFT-FT pair in the junction of iterations (in brackets) to reduce the computational time.

Depending on the NLS normalisation chosen and amplification technique used, there are different definitions of the operators  $\hat{\mathbf{L}}$  and  $\hat{\mathbf{N}}$  emerging to the computation. The transition to the spectral (frequency) domain requires proper sampling there, defined via the FT properties (see appendix A). The array  $\Omega$  is non-symmetrical array (including zero sample) with step  $\Delta\Omega = \frac{2\pi}{\Delta\vartheta n_t}$ , where  $n_t$  is number of signal samples and  $\Delta\vartheta$  is sampling step in temporal domain. The operators, acting in the exponents, are as follows:

$$\hat{\mathbf{D}} = -iC_{\text{disp}}\Omega^2, \quad \hat{\mathbf{N}} = iC_{\text{nonl}}\mathbf{f}^2. \quad (2.18)$$

The loss term from eq. (2.8) can be taken into account adding  $-C_{\text{loss}}$  to any of the operators, since the FT is linear operation.

The expressions for the coefficients values for different forms of NLS are provided in table 2.1.

To simulate EDFA amplification, one should take into account losses during the span and then locally add gain, introduced by the amplifier at the end of each span. To simulate Raman amplification, one should consider either constant or smoothly varying gain profile all through the span. Both amplification formats introduce ASE noise to the system.

## 2.4 Optical communications

Optical fibre systems form the backbone of the global telecommunication networks that underpin Internet, broadband communications and digital economy. It is hard to overstate the impact that fibre communications have made on the economy, society, and almost all aspects of our lives. Over the last century, the achievements of optical communications supported the growth of global economic, enabling the effective communication of various distance and rates scales.

The celebrated estimation is done by Shannon [127] for the maximum achievable capacity of the linear channel in the presence of the additive noise:

$$\text{capacity} = \text{bandwidth} \cdot \log_2(1 + \text{SNR}). \quad (2.19)$$

Note that this estimation has been done regardless of optical fibre, however in the linear limit (neglecting Kerr and other nonlinearities) it applies to fibre optics as well. This formula effectively means that for expanding the channel capacity one has to either exploit

the wider bandwidth or increase the signal-to-noise ratio (SNR). However, the optical channel model deviates from linear, especially in the high-SNR regime. Therefore, the nonlinearity severely bounds the capacity and does not allow the optical-fibre channel to achieve this anticipated data rates.

Subsequent discoveries of wavelength-division multiplexing, EDFA, coding and constellation shaping approaches allowed satisfying the demand in the information channel capacity. Current forecasts predict the exponential growth of the data traffic demand (by other economic sectors) [92]. The growth of the Internet-based activities and the rapid proliferation of bandwidth-hungry online services (cloud computing, big data, on-demand HD video streams, Internet of Things, etc.) have resulted in the ever-escalating pressure on the speed and quality of information flows interconnecting network participants. Hence, contemporary communication systems are facing an increasing challenge to provide more and more line capacity each year.

For long-haul communications, the fibre nonlinearity is considered as the most principal limiting component. The currently employed systems demonstrate the saturation, and then degradation of the achievable channel capacity (measured, e.g., by data rate or spectral efficiency (SE)). This saturation and degradation is observable for high SNR regimes and occur earlier for longer distances/higher data rates (see fig. 2.2). However, fibre-optics technologies, especially after the developing of efficient amplification formats and exploiting advanced coding approaches, are still the most efficient and powerful for large-scale transmission. Therefore, the compensation of nonlinear distortions is a crucial factor for the further development of the communication systems.



Figure 2.2: The achievable data capacity by optical fibre communication systems, as a function of link length and SNR (cited from [36]).

## 2.5 Nonlinearity mitigation techniques

Responding to the data traffic demand and its forecasting for the nearest future, various research groups are approaching the problem of an increase in transmission capacity from different sides. In particular, one of the approaches aims to avoid significant fibre re-installation costs via developing advanced signal processing methods which reduce the effects of the nonlinearity to the transmitted information. These approaches are jointly named in the literature as nonlinearity mitigation techniques.

One of the methods is the so-called optical phase conjugation (OPC). In this technique, the phase factor of the incident wave is inverted, swapping its sign. In NLS it can

be expressed by replacing the electrical field envelope  $A(t, z)$  with its complex conjugation  $A^*(t, z)$ . After propagating on the same distance, the accumulated effect of both chromatic dispersion and Kerr nonlinearity are inverted and compensated. This approach was first suggested in [149]. The further developing in [24, 35, 61, 114, 143] demonstrated the performance of this approach for the transmission system with different modulations and with distances up to 10,000 km [61]. However, for ideal compensation, it requires symmetrical power profile on both sides of conjugation point, what is unrealistic for long-distance EDFA-aided transmission systems. Also, the input-frequency bandwidth is essentially limited by the phase-matching requirement of the specific nonlinear mixing scheme.

Another well-developed approach to nonlinearity mitigation is the so-called digital backpropagation (DBP). In the contrast with OPC, this compensation is done by means of the signal processing in the digital domain, but not in optical methods. It reduces the achievable rates for this method. Practically, the method consists in solving the NLS model of the channel, including the specific gain profile, in the backward direction, starting from the receiver profile and getting the transmitted one [28, 56, 88, 120].

At the moment, DBP approach is considered now as the most powerful in full compensation of all deterministic impairments. It is naturally cannot directly deal with the noise appearing in the optical channel, and with the inter-channel interaction. There are some attempts to implement DBP using hardware application-specific integrated circuit (ASIC) technologies, aiming application for the properly real-time transmission line [112, 121], or to further reduce the numerical complexity of the algorithm [106]. The overall computational complexity of this method, assuming split-step Fourier method with FFT, see section 2.3, is  $O(N_z \cdot N_q \log_2 N_q)$ . The number of steps  $N_z$  is apparently crucial to get acceptable numerical accuracy, therefore the application of DBP is considered by the community as a computationally-expensive method.

The inverse Volterra series provided similar backward propagation of the received signal, using digital compensation algorithms. Presented in [113], it applied to the frequency domain image of the input signal and links it to the Fourier image of the output waveform. The correspondence is expressed in the sequence of integrals over the partial transfer functions. The method can be applied for generalised forms of NLS, including additional terms, for example, for Raman effect [48, 91, 111, 113].

This application of this method is based on the trade-off between the order of the series truncation, and desirable accuracy, determined by the parameters of the pulse. Increase of the number of integration in the sequence  $N_{\text{int}}$  dramatically influence the computational complexity as  $O(N_q^{N_{\text{int}}-1})$ .

Recent advances in developing of machine learning (ML) algorithms, especially significant lowering of the computational efforts due to parallelisation, is actively applied for needs of optical communications.

In general, ML methods are implemented for objects classification, linear regression, recognition, etc. In the application for optical communications, some ML techniques are applied for network routing [105, 150], channel monitoring and identification [20, 68, 105, 155], constellation equalisation [45, 155, 156], backpropagation [7, 69, 105, 128, 148] and others [39, 68, 69, 148].

My research during the PhD study included co-authoring several publications about application of ML methods for filtering and equalisation of the signal on the receiver (Rx) for NFT-based transmission [CC2, CC3].

## 2.6 Additional details for transmission simulations

The optical communication line consists of different compounds and the adjusting of the parameters is crucial for sufficient transmission. The numerical simulation of the optical

channel consists of three principal stages: modelling transmitter (Tx) side, fibre propagation and Rx side. The details of the simulation are given in the following subsections.

### 2.6.1 Modulation approaches

The (bit) data stream on the Tx needs to be converted to the waveform in the way, which ensures the optimised performance and the transmission quality. The approach to this in the NFT-based transmission is inherited from the conventional transmission, where also the waveform is to be created from the bunch of data, truncated from the data stream.

The modulation is usually depicted using the so-called *constellation diagram* and the set of possible complex numbers is named constellation [34, 36, 70]. The most simple modulation keying is on-off keying, where the light phase is not affected and only the presence or the absence of the pulse is modulated. The developments of coherent detection approached unlocked the opportunity to modulate both the amplitude and the phase. Mathematically, it means the extension to the set of complex numbers from the set of real numbers. The typical approach there are quantised modulations, i. e. quadrature amplitude or phase modulations. Each point of the constellation is corresponded to the specific binary number, reflected binary code, also known as Gray code, where successive values differ in one binary digit. The main constellation types with corresponding properties are listed in table 2.2.

### 2.6.2 Single-mode fibre parameters

In this research, I work within the NFT approach for communication, what hiddenly means that initial normalisation of NLS does not change during the signal propagation. As the normalisation is predefined by the fibre parameters,  $\alpha$ ,  $\beta_2$  and  $\gamma$ , along the optical channel, where NFT is applied, these parameters cannot be changed. In addition, one of the principal advantages of NFT-based communication is an unlocked the opportunity to use more intelligently the same, already embedded, fibre infrastructure. Due to these facts, the main results of this thesis are done under the assumption of usage of SMF. The results can be re-evaluated and generalised to other fibre properties, however, it is principally required to have the same fibre properties along the entire transmission line.

The SMF parameters and other relevant constants are presented in table 2.3 [4, 36, 80].

### 2.6.3 Transmission quality metrics

The performance of the transmission systems is expressed in the form of the achieved data rate at some propagation distance. Instead of data rate, one can use SE: if the data rate is a number of bits per unit time, measured in [bit/s], then the SE is a number of bits per unit time and unit bandwidth, i.e., measured in [bit/s/Hz].

For the designed system with the given data rate, the common metrics of the performance is quality-, or Q-factor, presented as a function of the signal power. It can be determined either from bit error rate (BER), or from error-vector magnitude (EVM). BER is a number of wrongly detected bits, normalised over a total number of bits, i.e.:

$$\text{BER} = \frac{\# \text{ wrong bits}}{\text{total } \# \text{ bits}}. \quad (2.20)$$

EVM is defined as an error vector norm normalised over a signal vector norm. Say, the signal is from the constellation set,  $c$ , so EVM is given as:

$$\text{EVM}^2 = \frac{\langle |c^{Rx} - c^{Tx}|^2 \rangle}{\langle |c^{Tx}|^2 \rangle}. \quad (2.21)$$

Table 2.2: Conventional modulation constellation structures.

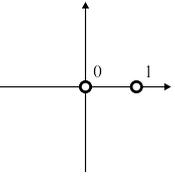
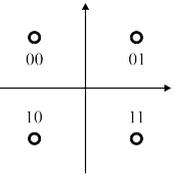
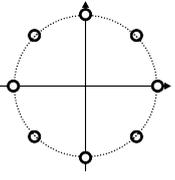
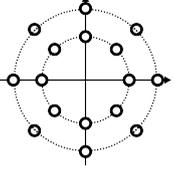
Constellation	Name	bits/symbol
	on-off keying (OOK)	1
 <p style="text-align: center;"><math>N = 4</math></p>	$N$ -quadrature amplitude modulation ( $N$ -QAM)	$\log_2 N$
 <p style="text-align: center;"><math>M = 8</math></p>	$M$ -phase-shift modulation ( $M$ -PSK)	$\log_2 M$
 <p style="text-align: center;"><math>M = 4, N = 2</math></p>	$M$ -PSK/ $N$ -ASK	$\log_2 MN$

Table 2.3: Single-mode fibre parameters in the vicinity of C-band.

Quantity	Name	Value (if possible) and unit
$\alpha$	attenuation coefficient	0.2 dB/km = $4.65 \cdot 10^{-5} \text{ m}^{-1}$
$\beta_2$	GVD coefficient	$-22 \cdot 10^{-27} \text{ s}^2 \cdot \text{m}^{-1}$
$\gamma$	Kerr nonlinear parameter	$1.27 \cdot 10^{-3} \text{ W}^{-1} \cdot \text{m}^{-1}$
$\lambda_0$	carrier wavelength	1550 nm
$f_0$	carrier frequency	$1.93 \cdot 10^{14} \text{ Hz}$
$\omega_0$	carrier circular frequency	$1.2 \cdot 10^{15} \text{ rad} \cdot \text{s}^{-1}$
$A_{\text{eff}}$	effective core area	$80 \mu\text{m}^2$
$n_2$	nonlinear correction to $n$ , eq. (2.4)	$2.6 \cdot 10^{-20} \text{ m}^2 \cdot \text{W}^{-1}$
$D$	dispersion parameter	$17 \text{ ps} \cdot \text{nm}^{-1} \cdot \text{km}^{-1}$
$S$	dispersion slope	$0.075 \text{ ps} \cdot \text{nm}^{-2} \cdot \text{km}^{-1}$

The Q-factor value can be found from both EVM or BER:

$$Q_{\text{BER}} = 20 \log_2 \sqrt{2} \text{erfc}^{-1} \text{BER}, \quad Q_{\text{EVM}} = -20 \log_{10} \text{EVM}. \quad (2.22)$$

The forward error correction technique defines the threshold value of BER at  $7 \cdot 10^{-3}$ , the lower bound for Q-factor is determined correspondingly.

The improvement in Q-factor usually acts as evidence of the room available for further increase in the data rate.

## 2.6.4 Noise consideration in the numerical simulations

It took plenty of efforts and time to get a correct and consistent understanding of the noise generation in numerical simulations. Below I provide a comprehensive description of the procedures.

The necessity to embed an amplification to the transmission scheme in order to compensate the fibre loss, introduces the dominating source of noise into the system. It is attributed to the ASE effect (see subsection 2.1.4).

To the best of my knowledge, in theoretical and numerical studies the noise is modelled via adding the additive white Gaussian noise (AWGN) term to the r.h.s. of the NLS of either form and normalisation [32, 36, 135]. However, it is not a universal approach, there is also an effective way to model the noise via band-limited Wiener process [41]. In this thesis, I follow a more widely used model of AWGN.

The main parameter, which defines the intensity of ASE is a noise power spectral density  $\Gamma$ , i.e. power of a unit of spectrum bandwidth of a unit length of the fibre. It appears in the model as a coefficient of the correlator of the NLS noise term  $N(t, z)$ :

$$\langle N(t, z) N^*(t', z') \rangle = \Gamma \delta(t - t') \delta(z - z'), \quad (2.23)$$

where  $\delta(\cdot)$  designates the Dirac delta function, eq. (2). This formula reflects the fact that the ASE noise is additive, white and Gaussian. The additional requirement here is that  $N(\cdot, \cdot)$  is zero-mean function.

For Raman and EDFA expressions for  $\Gamma$  are ideologically identical, with the only difference that for EDFA as a typical distance one uses span length  $L_s$ , when for Raman - the entire length of a fibre link [36]:

$$\Gamma_{\text{EDFA}} = n_{\text{sp}} h f_0 (e^{\alpha L_s} - 1) / L_s, \quad \Gamma_{\text{Raman}} = \alpha h f_0 K_T, \quad (2.24)$$

where  $h f_0$  is the emitted photons energy,  $n_{\text{sp}}$  is the spontaneous emission factor and  $K_T$  is photon occupancy factor, equal to 1.13 at the room temperature.

For SMF with the parameters, listed in table 2.3, the numerical value of  $\Gamma = 6.6 \cdot 10^{-24}$  W/Hz/m.

The noise term of NLS experience the same rescaling as all other terms during the normalisation. For normalisation (i), one can add to the NLS in the form (2.14) or (2.13) the item  $n(\tau, \zeta)$  given by:

$$n^{(i)}(\tau, \zeta) = \frac{\gamma^{1/2} T_0^3}{|\beta_2|^{3/2}} N(T_0 \tau, L_0 \zeta). \quad (2.25)$$

For the equivalently defined normalised noise spectral density as:

$$\langle n^{(i)}(\tau, \zeta) n^{(i)*}(\tau', \zeta') \rangle = \Gamma_0^{(i)} \delta(\tau - \tau') \delta(\zeta - \zeta'), \quad (2.26)$$

its intensity value is given by  $\Gamma_0^{(i)} = \frac{\gamma T_0^3}{|\beta_2|^2} \Gamma$ .

In the case of normalisation (ii), the expressions for the rescaled noise term and the normalised noise power spectral density are:

$$n^{(ii)}(\tau, \zeta) = \frac{2i\gamma^{1/2} T_0^3}{|\beta_2|^{3/2}} N(T_0 \tau, L_0 \zeta), \quad \Gamma_0^{(ii)} = \frac{2\gamma T_0^3}{|\beta_2|^2} \Gamma. \quad (2.27)$$

Note that according to the Dirac delta function properties, I use the following relation:  $\delta(kx) = \delta(x)/k$ .

## Chapter 3

# Inverse scattering transform

In this chapter, I introduce the full description of the procedures and notions behind the Inverse Scattering Transform (IST) technique, which is the main object of the research presented in the current thesis.

The formulae used in this chapter, are related to the normalised integrable NLS in the form of eq. (2.15):

$$\frac{\partial q}{\partial \zeta} = i \frac{\partial^2 q}{\partial \tau^2} + 2i|q|^2 q,$$

and  $\lambda$  here and below designates the variable in the Nonlinear Fourier (NF) domain and, in general case, complex-valued. During writing this chapter, I was mainly referring to the well-written review by Aktosun [6].

The IST was suggested as a way of solution of some specific evolutionary nonlinear partial differential equations, so-called exactly solvable, or integrable equations [2, 43, 78, 79]. Then, it was reintroduced under the name of Nonlinear Fourier Transform [53, 151] to highlight the analogue between the generation of the spectrum utilizing linear FT and the mapping to the NF domain with the creation of nonlinear spectrum.

Lately, it was discovered the direct connection between the existence of solitary solutions of the nonlinear partial differential equations and their integrability. The solitary solutions (or solitons) is a specific waveform which conserves its shape because of the interplay between nonlinear effects and effects which lead to the wave package spreading. First, they were observed experimentally on the canal between Edinburgh and Glasgow by John Scott Russell in 1834 [124]. Now, the most interesting integrable equations from the application point are Korteweg-de Vries (KdV) equation, which appears as in shallow water waves dynamics [71] or ion-acoustic waves in a plasma [153], sine Gordon equation, used for Josephson effect in semiconductors [13], and the last but not least NLS, which apart from the optical fibre application, is used as a model for deep water waves.

The IST was first discovered by Gardner, Greene, Kruskal, and Miura [43] where they showed the connection between the KdV equation and 1D Schrödinger equation, when the latter acts as a corresponding scattering problem for the IST. For the NLS, Zakharov and Shabat in 1972 introduced the corresponding IST [154].

The way of solving the nonlinear partial differential equation employing IST is schematically shown in fig. 3.1. Having initial value problem, i.e. the waveform to be launched into the fibre, one can use specific mapping, known as direct scattering transform or direct (forward) NFT, to get the spectral representation of the initial value. Then, the propagation of the scattering data in the NF domain is effectively linear, it is the phase shift similar to the effect of the chromatic dispersion in the linear spectral domain. Getting the spectral data of the integrable equation solution, one should perform the counterpart mapping, this step is literally called inverse scattering transform, or inverse Nonlinear Fourier Transform (INFT). This procedure returns the solution of the integrable equation

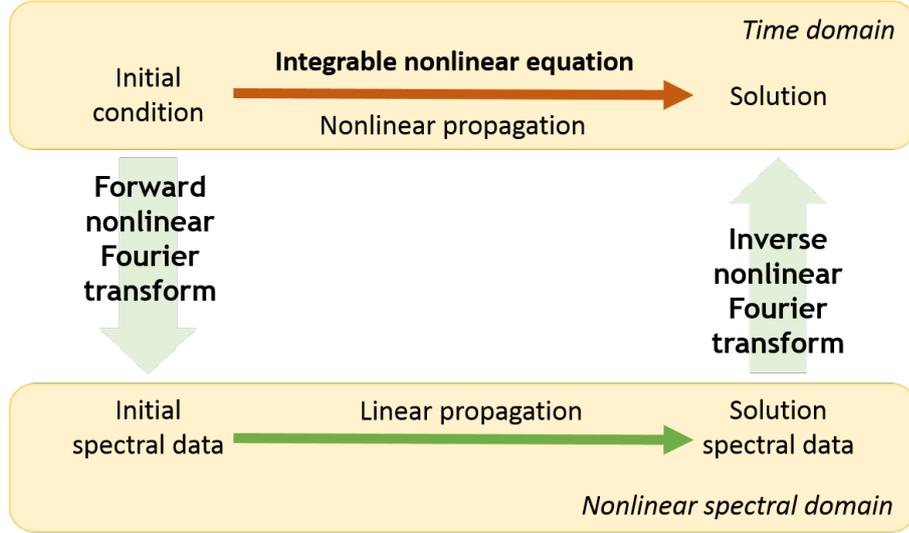


Figure 3.1: Schematic of the procedure of integrable equation solution using the IST approach.

at any point of interest. Although both inverse and forward NFT are mathematically complicated and computationally expensive procedures (despite their linearity), this mapping allows effectively simplify the nonlinear evolution according to the partial differential equation to linear evolution of scattering data, by the cost of performing two transform operations. Literally, the IST allows replacing the nonlinear evolution according to the given integrable equation by the row of three linear mappings, defined by this equation.

The same concept can be approached from the other side: starting from the given scattering problem (it is a broad definition, many equations can be interpreted as a scattering problem), it is possible to derive the corresponding evolutionary equation. One of the methods is the so-called Lax approach, [79], which uses the relations between operators of the scattering problems to get the integrable equation. One can define a spectral problem (with spectral parameter  $\lambda$ ) in 1+1 space  $(\tau, \zeta)$  for a two-element vector  $\psi(\tau, \zeta)$  as:

$$\hat{\mathbf{L}}\psi = \lambda\psi, \quad (3.1)$$

where  $\hat{\mathbf{L}}$  is a differential operator, it is possible to find a corresponding operator  $\hat{\mathbf{A}}$  (generally speaking, the hierarchy of operators  $\mathbf{A}$ , NLS is a part of this hierarchy for the operator  $\mathbf{A}$  of 2nd order). The pair of operators  $\hat{\mathbf{L}}$  and  $\hat{\mathbf{A}}$  is called *Lax pair* and they satisfy the following relations:

$$\begin{aligned} &\text{the spectral parameter is constant } \frac{\partial \lambda}{\partial \zeta} = 0, \\ &\text{transform } \psi \mapsto \frac{\partial \psi}{\partial \zeta} - \hat{\mathbf{A}}\psi \text{ still satisfies the spectral problem (3.1),} \\ &\text{the operator } \frac{\partial \hat{\mathbf{L}}}{\partial \zeta} + \hat{\mathbf{L}}\hat{\mathbf{A}} - \hat{\mathbf{A}}\hat{\mathbf{L}} \text{ is purely multiplicative.} \end{aligned} \quad (3.2)$$

Moreover, the third condition is more specific, joining all requirement together, one can get:

$$\frac{\partial \hat{\mathbf{L}}}{\partial \zeta} + \hat{\mathbf{L}}\hat{\mathbf{A}} - \hat{\mathbf{A}}\hat{\mathbf{L}} = 0, \quad (3.3)$$

which is called a *compatibility condition* and the specific form of it gives literally a desired evolutionary partial differential equation. In particular, for NLS, the explicit expressions

for the Lax pair operators  $\hat{\mathbf{L}}$  and  $\hat{\mathbf{A}}$  are as follows:

$$\hat{\mathbf{L}} = \begin{pmatrix} i\frac{\partial}{\partial\tau} & -iq \\ -iq^* & -i\frac{\partial}{\partial\tau} \end{pmatrix}, \quad \hat{\mathbf{A}} = \begin{pmatrix} 2i\frac{\partial^2}{\partial\tau^2} + i|q|^2 & -2iq\frac{\partial}{\partial\tau} - i\frac{\partial q}{\partial\tau} \\ -2iq^*\frac{\partial}{\partial\tau} - i\frac{\partial q^*}{\partial\tau} & -2i\frac{\partial^2}{\partial\tau^2} - i|q|^2 \end{pmatrix}. \quad (3.4)$$

The compatibility condition (3.3) gives the NLS for  $q(\tau, \zeta)$  in the (12) entry, NLS for  $q^*(\tau, \zeta)$  in the (21) entry and exact zero in (11) and (22) entries.

The operator  $\hat{\mathbf{L}}$  defines the scattering problem, as it was first derived by Zakharov and Shabat in [154], it is known as the Zakharov-Shabat problem (ZSP), commonly written as ordinary differential equations system as:

$$\frac{\partial\psi}{\partial\tau} = \begin{pmatrix} -i\lambda & q(\tau, \zeta) \\ -q(\tau, \zeta)^* & i\lambda \end{pmatrix} \psi. \quad (3.5)$$

Note that the form of the ZSP does not depend on specific coefficients on NLS, however it influences the evolution of the associated NF spectral data, defined by the operator  $\hat{\mathbf{A}}$ .

There is an alternative way of deriving the nonlinear partial differential equation related to the given scattering problem, suggested by Ablowitz, Kaup, Newell, and Segur in [1], now known as Ablowitz, Kaup, Newell, Segur (AKNS) method. Expressing ZSP as an ordinary differential problem for two-element vector:

$$\frac{\partial\psi}{\partial\tau} = \hat{\mathbf{T}}\psi, \quad \text{so } \hat{\mathbf{T}} = \begin{pmatrix} -i\lambda & q(\tau, \zeta) \\ -q(\tau, \zeta)^* & i\lambda \end{pmatrix}, \quad (3.6)$$

one need to find the corresponding operator  $\hat{\mathbf{Z}}$ , satisfying the following set of conditions

$$\begin{aligned} &\text{the spectral parameter is constant } \frac{\partial\lambda}{\partial\zeta} = 0, \\ &\text{transform } \psi \mapsto \frac{\partial\psi}{\partial\zeta} - \hat{\mathbf{Z}}\psi \text{ still satisfies the spectral problem (3.6),} \\ &\text{the operator } \frac{\partial\hat{\mathbf{T}}}{\partial\zeta} - \frac{\partial\hat{\mathbf{Z}}}{\partial\tau} + \hat{\mathbf{T}}\hat{\mathbf{Z}} - \hat{\mathbf{Z}}\hat{\mathbf{T}} = 0. \text{ (similarly to eq. (3.3)).} \end{aligned} \quad (3.7)$$

For NLS, the counterpart operator has the following form:

$$\hat{\mathbf{Z}} = \begin{pmatrix} i|q(\tau, \zeta)|^2 - 2i\lambda^2 & 2\lambda q(\tau, \zeta) + i\frac{\partial q(\tau, \zeta)}{\partial\tau} \\ i\frac{\partial q^*(\tau, \zeta)}{\partial\tau} - 2\lambda q^*(\tau, \zeta) & 2i\lambda^2 - i|q(\tau, \zeta)|^2 \end{pmatrix}. \quad (3.8)$$

Likely to the Lax approach, the comparability condition yields NLS, the operator  $\hat{\mathbf{T}}$  defines the problem for direct scattering, i.e. ZSP, and the operator  $\hat{\mathbf{Z}}$  determines the evolution of the scattering data. In the contrast with Lax approach, where the operator  $\hat{\mathbf{A}}$  is independent on NF spectral parameters  $\lambda$ , in AKNS method both operators depend on  $\lambda$ .

The IST method described above, is applicable for the initial value problem of any integrable equation. Since NLS belongs to this class, so one can utilize this technique for any purposes. However, for the complete statement of the initial value problem, we also need to specify the boundary conditions, so some constraints for  $q(\tau, \zeta)$ , being true for any  $\zeta$ . The following sections describe the remaining details of IST approach for two cases interesting for applications: vanishing boundary conditions, when function  $q(\tau, \zeta)$  decays sufficiently fast at  $\tau \rightarrow \pm\infty$ ; and periodic boundary condition, when we require strict periodicity for  $q(\tau, \zeta)$  with some positive period, maintaining constant during the profile evolution.

From the practical point of view, there is no preference in choosing one of the cases: effectively, the optical fibre introduces different medium-light interactions and NLS is only

a leading-order model of the signal propagation in the fibre. In particular, the optical fibre does not have any symmetry which can allow assuming some specific boundary condition, from the physics point of view, the fibre solves NLS for the infinite continuous signal. The abovestated reasoning means that the choice of boundary conditions should be done from the practical reasons, having in mind the application for communications. More about the difference between vanishing and periodic boundary conditions in NFT one can find in section 6.1.

## 3.1 Vanishing boundary conditions

### 3.1.1 Local definitions

The full description of the IST requires to specify the boundary condition. The so-called vanishing boundary condition NFT is more common in literature because it generalises the well-known idea of solitonic structures in integrable equations. Moreover, the idea of utilizing solitonic modes of communication purposes was suggested in [54], and then further developed in [52]. In general, the vanishing boundary condition case allows more degrees of freedom and more flexibility to monitor the signal parameters.

From the mathematical point of view, apart from the obvious condition for the limit at infinity:

$$\lim_{\tau \rightarrow \pm\infty} q(\tau, \zeta) = 0, \quad \forall \zeta, \quad (3.9)$$

it is essential to add the so-called limited  $L_1$ -norm condition:

$$\int_{-\infty}^{\infty} |q(\tau, \zeta)| d\tau < \infty, \quad \forall \zeta. \quad (3.10)$$

The spectral characteristics of the initial signal consist of three principal parts: (i) *continuous data*, defined in terms of the spectral (scattering) functions of real spectral variable, (ii) *solitonic eigenvalues*, distinguished discrete values of spectral parameter, and (iii) respective *norming constants* (also called spectral amplitudes) associated with the eigenvalues. Note that either continuous spectral data or discrete spectrum, consisting of pairs of an eigenvalue and corresponding spectral amplitude, can be absent for the specific potentials. For example, discrete-spectrum-only profiles are known as ‘reflection-less’ potentials from the point of the scattering problem. Accurate estimation of both parts of discrete NF data is critically important as they contain information about the bound states (solitons). It is worthwhile to note that the two last components vanish in the linear limit (i.e. for the pulses with the energy tending to zero), when NFT reduces to the conventional FT [117, 135]. Because of this reason, the spectral parameter is commonly called *nonlinear frequency* as it has the same meaning in the linear limit (to be precise,  $\lambda = -2\omega$  [117]).

For some cases, one can also require the sufficient decay rate  $d$  of the potential  $q(\tau, \zeta)$  for fixed value of  $\zeta$  defined as:

$$|q(\tau, \zeta)| < De^{-d|\tau|}, \quad \text{for all } \tau \in (-\infty, \infty), \quad (3.11)$$

which influences the analytical behaviours of the scattering functions in the complex plane of the spectral parameter  $\lambda$ .

In order to find the full set of the scattering data, we use eq. (3.5), substituting there the profile  $q(\tau, \zeta)$  at some cross-section point along the fibre  $\zeta$ . Therefore, we can, keeping the generality, give up the dependency on  $\zeta$  for a while, since the definition and procedures of the spectral NF mapping of the signal does not depend on the location in the fibre. ZSP here can be interpreted as a scattering problem for the coupled waves, approaching the scattering potential  $q(\tau)$  from the left and/or from the right, i.e from  $\tau \rightarrow \pm\infty$ .

Using notions of left-to-right wavefunctions for  $\psi$ , the so-called Jost functions, one can fix asymptotes at left infinity (valid for the spectral parameter  $\lambda$  from the closed upper half-plane  $\lambda \in \overline{\mathbb{C}_+}$ ):

$$\psi^{lr}(\tau, \lambda) \rightarrow \begin{pmatrix} \exp(-i\lambda\tau) \\ 0 \end{pmatrix} \quad \text{as } \tau \rightarrow -\infty, \quad (3.12)$$

or at right infinity:

$$\psi^{rl}(\tau, \lambda) \rightarrow \begin{pmatrix} 0 \\ \exp(i\lambda\tau) \end{pmatrix} \quad \text{as } \tau \rightarrow \infty. \quad (3.13)$$

In the following text, I mainly stick to the definitions via left-to-right waves, i.e. when the wavefunctions are defined by their asymptotes at  $\tau \rightarrow -\infty$ , eq. (3.12). For all  $\lambda \in \overline{\mathbb{C}_-}$  (lower complex half-plane), we additionally define  $\hat{\psi}^{rl}(\tau, \lambda) = (\psi_2^{rl*}(\tau, \lambda^*), -\psi_1^{rl*}(\tau, \lambda^*))^T$ , where  $\psi_{1,2}$  are entries of the vector  $\psi$ .

The left and right coupled wavefunctions are related via the scattering matrix, having the following form:

$$\psi^{lr}(\tau, \lambda) = \psi^{rl}(\tau, \lambda) \mathbf{S}(\lambda), \quad \mathbf{S}(\lambda) = \begin{pmatrix} a(\lambda) & -b^*(\lambda^*) \\ b(\lambda) & a^*(\lambda^*) \end{pmatrix}, \quad (3.14)$$

with  $\det \mathbf{S}=1$ .

The continuous scattering data are defined via functions  $a(\lambda)$  and  $b(\lambda)$  for  $\lambda \in \mathbb{R}$  by

$$\psi^{lr}(\tau, \lambda) = \hat{\psi}^{rl}(\tau, \lambda) a(\lambda) + \psi^{rl}(\tau, \lambda) b(\lambda), \quad \lambda \in \mathbb{R} \quad (3.15)$$

with

$$a(\lambda) = \det(\psi^{lr}(\tau, \lambda), \psi^{rl}(\tau, \lambda)) \quad \text{and} \quad b(\lambda) = \det(\hat{\psi}^{rl}(\tau, \lambda), \psi^{lr}(\tau, \lambda)). \quad (3.16)$$

In view of (3.12), they can also be expressed by

$$a(\lambda) = \lim_{\tau \rightarrow +\infty} \psi_1(\tau, \lambda) e^{i\lambda\tau}, \quad b(\lambda) = \lim_{\tau \rightarrow +\infty} \psi_2(\tau, \lambda) e^{-i\lambda\tau}. \quad (3.17)$$

They can also be characterized in terms of a single function, the (right or left) reflection coefficient, commonly called the continuous NF spectrum, related to dispersive waves within the potential:

$$r(\lambda) = \frac{b(\lambda)}{a(\lambda)}, \quad l(\lambda) = \frac{b^*(\lambda^*)}{a(\lambda)} \quad (3.18)$$

The solitonic degrees of freedom are associated with the discrete spectral data consisting of the set of complex-valued eigenvalues  $\{\lambda_j\}$  of eq. (3.5), that have positive imaginary parts (the set is finite due to conditions from eqs. (3.10-3.11), and the set can be empty), together with complex-valued norming constants (spectral amplitudes)  $\{c_j\}$ . Since  $\psi^{lr}(\tau \rightarrow -\infty, \lambda) \rightarrow 0$  and  $\psi^{rl}(\tau \rightarrow \infty, \lambda) \rightarrow 0$ , for any  $\lambda$  with  $\Im\lambda > 0$ , it follows that an eigenvalue  $\lambda_j$  is characterized by the linear dependence of  $\psi^{lr}(\tau, \lambda)$  and  $\psi^{rl}(\tau, \lambda) \rightarrow 0$ , i. e. by the existence of a non-zero constant  $b_j \in \mathbb{C}$  such that

$$\psi^{lr}(\tau, \lambda_j) = \psi^{rl}(\tau, \lambda_j) b_j. \quad (3.19)$$

On the other hand, as shown in eq. (3.16), the eigenvalues can be equivalently characterized as zeros of  $a(\xi)$  in the upper half-plane:  $a(\lambda_j) = 0$ . Generally, for potentials satisfying only eq. (3.10), the constants  $b_j$  are independent of the continuous scattering functions, but if  $\hat{\psi}^{rl}(\tau, \lambda_j)$  (and thus  $b(\lambda)$ ) also admits analytic continuation into the upper half-plane (or at least in the domain  $0 < \Im\lambda < d$ , preferably for  $d > \max \Im\lambda_j$ ), the parameters  $\{b_j\}$  represent the values of  $b(\lambda)$  evaluated at  $\lambda = \lambda_j$  (cf. (3.15)):

$$b_j = b(\lambda_j). \quad (3.20)$$

In this case, the norming constants can be defined as the residues of the reflection coefficient ( $r(\lambda)$  or  $l(\lambda)$ ) at its poles:

$$c_j = \text{Res}[r(\lambda)]|_{\lambda=\lambda_j} = \frac{b(\lambda_j)}{a'(\lambda_j)}, \quad d_j = \text{Res}[l(\lambda)]|_{\lambda=\lambda_j} = \frac{b^*(\lambda_j^*)}{a'(\lambda_j)}. \quad (3.21)$$

Particularly, for finitely supported  $q(\tau)$  (which is the case of the computational statement of the problem),  $b(\lambda)$  is analytic in the whole plane and thus definition (3.20) holds. In general case,  $b(\lambda)$  is analytic in the band in the  $\lambda$ -plane  $\Im\lambda < d$ , where  $d$  is the decay rate from eq. (3.11).

The NF image of the vanishing signal can be visualised in the form of figure 3.2.

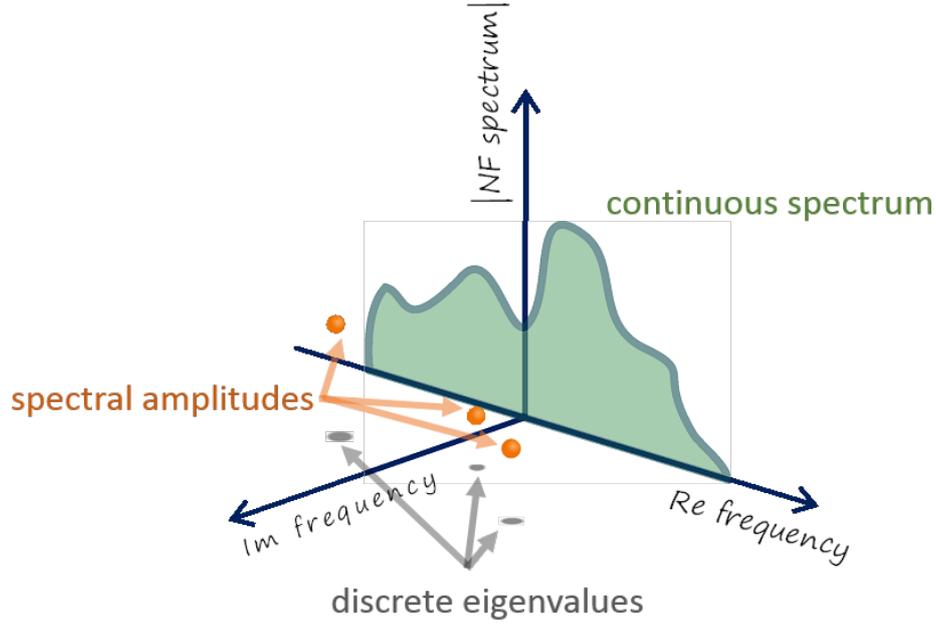


Figure 3.2: Visual representation of the NF image for the vanishing boundary condition case. It consists of a continuous spectrum as a waveform defined on the real axis of  $\lambda$ , say  $r(\lambda)$ , solitonic eigenvalues  $\lambda_j$  in the upper complex half-plane, and corresponding spectral amplitudes  $c_j$ .

Here I have to note that the scattering data are generally coupled objects, also possessing certain mathematical properties, which adds constraints to the range of values they can take. For example, there is the so-called *nonlinear Parseval's identity* [151], which makes a link between the energies (some profiles' weight measures) in the temporal and in NF domains, for the following expressions for continuous spectrum energy:

$$\begin{aligned} \varepsilon_{\text{cont}} &= -\frac{1}{\pi} \int_{-\infty}^{\infty} \log |a(\lambda)|^2 d\lambda \\ &= -\frac{1}{\pi} \int_{-\infty}^{\infty} \log (1 - |b(\lambda)|^2) d\lambda \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} \log(1 + |r(\lambda)|^2) d\lambda, \end{aligned} \quad (3.22)$$

and the discrete NF spectrum energy:

$$\varepsilon_{\text{disc}} = 4 \sum_j \Im\lambda_j. \quad (3.23)$$

The identity links the energy in NF domain with the signal's one, given by:

$$\varepsilon_q = \int_{-\infty}^{\infty} |q(\tau)|^2 d\tau, \quad \varepsilon_q = \varepsilon_{\text{cont}} + \varepsilon_{\text{disc}}. \quad (3.24)$$

The NF continuous spectrum's energy definition, eq. (3.22), uses the fact that because of the coupled waves  $\psi_{1,2}$  total energy conserves, the auxiliary scattering data satisfy:

$$a(\lambda)a^*(\lambda^*) + b(\lambda)b^*(\lambda^*) = 1. \quad (3.25)$$

The energy balance is actively used to predict the presence of the discrete eigenvalues, to estimate maximum eigenvalue's imaginary part, and to validate the performance of numerical routines, like it is done in [J2] and [12, 19, 151].

As we can conclude from the constraints above, different components of NF spectrum showcase the dependency on each other. So, one cannot define all of them independently to generate the desired solution of NLS. Therefore, we have to specify a *minimal sufficient set* of NF spectrum objects, which uniquely determines the corresponding profile  $q(\tau)$ . There is some freedom to include some or other objects to this set, depending on the designed transmission system and other reasoning.

First, there is an equivalent liability of left and right scattering data. Despite the right scattering data are more common in optical communication application [11, 140], in the application for fibre Bragg grating the left scattering data are used [38]. If one sticks to a certain (let it be right) set, there is still some freedom to play with.

The most straightforward way is to use  $r(\lambda)$  scattering function, it uniquely determines the continuous data. Moreover, it provides some constraints to the discrete spectrum as well, because the eigenvalues should coincide with the poles of  $r(\lambda)$ ,  $\lambda \in \mathbb{R}$  analytical continuation. Or one can start with the function  $b(\lambda)$ , as it is done in b-modulation approach. The waveform for  $b(\lambda)$ ,  $\lambda \in \mathbb{R}$  uniquely determines the waveform for  $a(\lambda)$ . For the discrete spectrum, there are two alternatives to use either norming constants  $c_j = b(\lambda_j)/a'(\lambda_j)$ , or the  $b_j = b(\lambda_j)$  to specify minimal sufficient information about the discrete scattering data. More about the usage of different nonlinear spectrum components for modulation the reader can find in chapter 6.

### 3.1.2 Alternative determination of the scattering data via ZSP for enveloping functions

It can be useful to rewrite ZSP (3.5) for the wave envelope functions for real-valued  $\lambda$   $\chi_{1,2}$  defined through the relations:

$${}_1(\tau, \lambda) = \chi_1(\tau, \lambda)e^{-i\lambda\tau}, \quad {}_2(\tau, \lambda) = \chi_2(\tau, \lambda)e^{i\lambda\tau}. \quad (3.26)$$

Then the ZSP for the envelope vector  $\mathbf{X}(\tau, \lambda) = (\chi_1(\tau, \lambda), \chi_2(\tau, \lambda))^T$  becomes

$$\frac{\partial}{\partial \tau} \mathbf{X}(\tau, \lambda) = \begin{pmatrix} 0 & qe^{2i\lambda\tau} \\ -q^*e^{-2i\lambda\tau} & 0 \end{pmatrix} \mathbf{X}(\tau, \lambda). \quad (3.27)$$

In terms of  $X$ , the spectral coefficients are given by

$$a(\lambda) = \lim_{\tau \rightarrow +\infty} \chi_1(\tau, \lambda), \quad b(\lambda) = \lim_{\tau \rightarrow +\infty} \chi_2(\tau, \lambda). \quad (3.28)$$

Since the initial conditions for  $\mathbf{X}(\tau, \lambda)$  in eq. (3.27) do not involve exponentials, which can behave rather oscillatory and make the numerical computation more problematic:

$$(\chi_1(\tau, \lambda), \chi_2(\tau, \lambda))^T \rightarrow (1, 0)^T \quad \text{as } \tau \rightarrow -\infty, \quad (3.29)$$

the definition of spectral coefficients via  $\mathbf{X}(\tau, \lambda)$  turns to be convenient for some numerical methods, derived in this research and described below.

In the computation of  $b_j$  for the norming constants or just for description of the scattering data (because of the linear dependency,  $b_j$  is sufficient to be used instead norming constants  $c_j$ ), there has been reported an issue of numerical divergence of algorithms used, see [10, 51]. To avoid this problem, the so-called bi-directional approach is used [J2], [51], when instead of finding the coupled waves values  $\psi$  at negative infinity and compare them with amplitudes at positive infinity of parameter  $\tau$ , one can find the ratio of waves amplitudes at the  $\tau = 0$ , where exponentials in the wave asymptotes do not reveal any fast-growing behaviour.

Using the wave function envelopes for both  $\psi^{lr}(\tau, \lambda)$  and  $\psi^{rl}(\tau, \lambda)$ , defined as

$$\mathbf{G}(\tau, \lambda) = \psi^{lr}(\tau, \lambda)e^{i\lambda\tau}, \quad \mathbf{H}(\tau, \lambda) = \psi^{rl}(\tau, \lambda)e^{-i\lambda\tau}, \quad (3.30)$$

we can evaluate the coupled systems for both vectors  $\mathbf{G}(\tau)$  and  $\mathbf{H}(\tau)$ :

$$\frac{\partial}{\partial\tau}\mathbf{G}(\tau, \lambda) = \begin{pmatrix} 0 & q(\tau) \\ -q(\tau)^* & 2i\lambda \end{pmatrix} \mathbf{G}(\tau, \lambda), \quad (3.31)$$

$$\frac{\partial}{\partial\tau}\mathbf{H}(\tau, \lambda) = \begin{pmatrix} -2i\lambda & q(\tau) \\ -q(\tau)^* & 0 \end{pmatrix} \mathbf{H}(\tau, \lambda). \quad (3.32)$$

To use these equations for the computation, the initial value of the vector  $\mathbf{H}(\infty, \lambda) = (0, 1)^T$  evolves from  $\tau \rightarrow \infty$  towards  $\tau = 0$ , whilst the vector  $\mathbf{G}(-\infty, \lambda) = (1, 0)^T$  evolves from  $\tau \rightarrow -\infty$  to  $\tau = 0$ . At the point  $\tau = 0$ , the desired quantity  $b_j$  can be obtained from the following relation (cf. (3.19)):

$$\mathbf{G}(\tau = 0, \lambda_j) = \mathbf{H}(\tau = 0, \lambda_j)b_j. \quad (3.33)$$

Here I have to emphasize that the presented way of definition and computation of  $b$  is appropriate only for the discrete eigenvalues.

### 3.1.3 Relations between soliton parameters and spectral data

The discrete spectral data are naturally related to the parameters of a certain bound state (soliton). The soliton is fully described by four parameters: amplitude  $A_j$ , frequency (also called velocity)  $\Omega_j$ , (centre-of-mass) position  $T_j$  and phase  $\varphi_j$ . The fundamental single soliton solution of the normalised NLS, eq. (2.15), has the following form [78]:

$$q(\tau, \zeta = 0) = A \operatorname{sech}[A(\tau - T)]e^{-i\Omega\tau + i\varphi}. \quad (3.34)$$

Soliton position and phase change their values during the evolution in  $\zeta$  according to the following expressions:

$$T(\zeta) = T(0) - \Omega\zeta, \quad \varphi(\zeta) = \varphi(0) + \frac{\zeta}{2}(A^2 - \Omega^2), \quad (3.35)$$

while the initial amplitude and frequency remain unchanged.

To find the relation between soliton parameters and the scattering data (here I use right definition  $c_j$ , however, one can derive the similar full system of relations, using left scattering data  $d_j$ ). The solution of ZSP for the single-soliton profile (3.34), from [78, ex. to 3.9]:

$$\begin{aligned} \psi_1(\tau) &= \frac{e^{-i\lambda\tau}}{\lambda - \Omega/2 + iA/2} \left( \lambda - \frac{\Omega}{2} - i\frac{A}{2} \tanh[A(\tau - T)] \right), \\ \psi_2(\tau) &= \frac{e^{-i\lambda\tau}}{\lambda - \Omega/2 + iA/2} \left( -\frac{iA}{2} \right) e^{i\Omega\tau - i\varphi} \operatorname{sech}[A(\tau - T)]. \end{aligned} \quad (3.36)$$

It leads to the following expression for the scattering coefficient  $b(\lambda)$  from (3.17) for the only eigenvalue  $\lambda_1 = \Omega/2 + iA/2$ , where  $a(\lambda_1) = 0$ :

$$b(\lambda_1) = -e^{-i\varphi} e^{AT}. \quad (3.37)$$

For  $a'(\lambda_1)$  which is needed for spectrum amplitude evaluation  $c_j$ , eq. (3.21), the expression is

$$a'(\lambda_1) = \frac{-4iA}{(A - 2i\lambda + i\Omega)^2} \xrightarrow{\lambda \rightarrow \lambda_1} -\frac{i}{A}, \quad (3.38)$$

resulting in the following formula for spectral amplitude:

$$c(\lambda_1) = \frac{b(\lambda)}{a'(\lambda)} \Big|_{\lambda=\lambda_1} = -iAe^{AT} e^{-i\varphi}. \quad (3.39)$$

Therefore, we obtain the following expressions for soliton parameters from NFT discrete spectrum:

$$A = 2\Im\lambda_1, \quad \Omega = 2\Re\lambda_1, \quad T = \frac{\log(|c(\lambda_1)|/A)}{A}, \quad \varphi = -\arg c(\lambda_1) - \pi/2. \quad (3.40)$$

For multisoliton profile one can use the same definitions, however, their physical meaning is unlikely to be so straightforward. Using the relation between spectral amplitudes  $c(\lambda_j)$  and  $T_j$ , one can estimate the location of the solitonic modes and adjust them to fit the dedicated time window.

### 3.1.4 Evolution of scattering data

The second operator from Lax pair  $\mathbf{A}$  (in Lax formalism), or from AKNS pair  $\mathbf{Z}$ , determines the evolution of the scattering data whilst the signal propagates according to NLS. Note that the formulae in this subsection are related explicitly to the normalisation of NLS, chosen for this research, i.e. eq. (2.15). Recall that signs in this equation are related to fibre's anomalous dispersion, where the existence of the solitonic modes is possible.

The fact that IST provides the effective linearization of the nonlinear channel means that the set of scattering data (either left or right) develops linearly whilst the signal propagates according to the integrable equation. By the constrains, which are used to derive the Lax pairs and AKNS pairs, eqs. (3.2) and (3.7), respectively, the discrete eigenvalues, if any, remain constant, i.e.

$$\lambda_j(\zeta) = \lambda_j(\zeta = 0). \quad (3.41)$$

The left set of the scattering data,  $l(\lambda)$  and  $d(\lambda_j)$  changes as [2, 78, 151]:

$$l(\lambda, \zeta) = l(\lambda, \zeta = 0) \exp(4i\lambda^2\zeta), \quad d(\lambda_j, \zeta) = d(\lambda_j, \zeta = 0) \exp(4i\lambda_j^2\zeta), \quad (3.42)$$

when the right scattering data satisfies:

$$r(\lambda, \zeta) = r(\lambda, \zeta = 0) \exp(-4i\lambda^2\zeta), \quad c(\lambda_j, \zeta) = c(\lambda_j, \zeta = 0) \exp(-4i\lambda_j^2\zeta). \quad (3.43)$$

In fibre optics application, for large distances, i.e. large value of the (normalised)  $\zeta$ , the exponential starts being highly oscillating, so it requires fine sampling to resolve this functional behaviour. Also, the wide effective NF bandwidth, i.e. occupation of far values of the spectral parameter  $\lambda$ , adds this oscillatory behaviour as well.

### 3.1.5 Discussion

To sum up, the vanishing boundary condition NFT is the most studied case from a mathematical perspective. It appears as a more transparent generalisation of linear spectrum, apart from the solitonic (discrete spectrum) component, which is a unique property of NFT processing. When the discrete modes responsible for the solitary waves, which maintain while signal propagation in the fibre, the continuous spectrum represents the dissipative waves. However, both components experience linear evolution according to the rule, similar to the dispersion effect in a linear channel. This and other similarities are convenient for the development of efficient communication systems.

Since the full NF spectrum picture consists of both continuous waveforms and discrete part, it provides a rich and flexible opportunity to modulate information in. In the case of the continuous spectrum, the amount of information to be modulated in it is limited only by the sampling step (in term of variable  $\lambda$ ). Atop it, the discrete-spectrum components also can be involved as data carriers. However, the design of NFT-based transmission system, one has to consider other practical advantages or drawbacks. For example, solutions modes usually include decaying tails which leak the informative part of the signal out of the dedicated time window. Additional, the numerical calculation of a discrete spectrum is less accurate (than for the continuous one) and it accumulates the numerical error, which may lead to the overall degradation of the demodulation accuracy.

## 3.2 Periodic boundary conditions

### 3.2.1 Local definitions

In the current research, I mainly base on the so-called Riemann-Hilbert problem approach: for INFT one can define the specific Riemann-Hilbert problem (RHP) and, solving it, find the periodic solution of NLS. In this approach, one should start from the specific structure of the NF spectral image. In general, RHP method allows the generation of various NLS solutions, however, the most straightforward way to get a periodic solution is to start with the so-called *finite-band* image in NF domain. This means that the profile can be parametrized by the finite number of domains (contours/bands/arcs) in the complex plane of the spectral parameter  $\lambda$ , or equivalently, a finite number of gaps between them.

In general, the solution, given by the RHP, is quasi-periodic, but not periodic, so it effectively consists of a finite number of modes which frequencies are determined by the NF spectrum. Adding more constrains to the picture in  $\lambda$  complex plane, the profile can be explicitly periodic when the frequencies are commensurable.

In the finite-band case, the image in NF domain is determined by the finite number of *main spectrum* points (complex numbers), and the contours are created by the connection of these points with their complex conjugated. Also the full representation of the periodic NLS solution in NF domain includes *auxiliary spectrum* part. The main difference between two part of the spectrum is that auxiliary one changes with  $\zeta$ , i.e. during the evolution of the initial profile according to NLS, whilst the main one maintains constant, like the discrete eigenvalues in the vanishing case. In RHP approach, the full picture of the finite-band solution requires to complement each contour in the complex plane with the real-value phase, chosen from the range  $[0, 2\pi]$ .

For the not-Hermitian spectral problem, eq. (3.4), the RHP formalism is less developed and more detached from applications, so there are no specified terms for some involved objects in NF domain. In a more studied case of Hermitian spectral problem, which corresponds for the *defocusing* NLS, or in the application for KdV equation [71, 100], the NF domain image consists of main spectrum points, located on real axis of spectral parameter  $\lambda$ . The main spectrum points are connected with arcs (cuts), known as *spectral*

*bands*, and intervals between them are known as *gaps*. Another feature of gaps is that the auxiliary spectrum points are localised on them. Therefore, in defocusing NLS there are two possible terms to characterise the type of solutions, either finite-band or finite-gap (meaning the finite number of one leads to the finite number of another). For the finite number of bands/gaps all gaps are finite objects when two bands include  $\pm\infty$  points, so, such solutions are more commonly known in the literature as finite-gap solutions.

Transiting to not-Hermitian ZSP, so to the focusing NLS, shifts the NF picture to the entire complex plane of  $\lambda$ . This picture still includes the arcs, connecting main spectrum points ( $\lambda$  and  $\lambda^*$  in this case), however, the picture is no longer on the real axis only, so the notion of gaps is losing its sense. Alternatively, we may treat a gap as a locus of auxiliary spectrum points, but there is a lack of evidence to specifically identify any regions in the complex plane, surrounding cuts, as a possible domain for auxiliary spectrum. Therefore, for the kind of solutions, studied here, and adapted to be periodic, the more accurate and appropriate term is finite-band solutions.

In this section, I list the notions and procedures of direct NFT for the periodic case. However, the reader must understand that this direct NFT is adjusted to the case when inverse NFT is performed employing RHP. The main spectrum of  $\mathcal{N}$ -genus NLS solution is defined as  $\mathcal{N} + 1$  complex points (bands) in the  $\lambda$ -plane  $\{\lambda_j\}_{j=0}^{\mathcal{N}}$ . The finite-genus solutions are generally quasi-periodic, but the periodicity can be ensured by imposing some additional constraints on the main spectrum. For example, using adjustment techniques, e.g., as in subsection 6.5.4, one can make an explicitly periodic profile. Main spectrum points define the Riemann surface function  $w(\lambda)$  as:

$$w^2(\lambda) = \prod_{j=0}^{\mathcal{N}} (\lambda - \lambda_j)(\lambda - \lambda_j^*), \quad (3.44)$$

choosing the branch of two-value function where asymptotically  $w(\lambda \rightarrow \infty) \sim \lambda^{\mathcal{N}+1}$ . The contours  $\Gamma_j = (\lambda_j, \lambda_j^*)$  connect the main spectrum points with their complex conjugate. Additionally, on each contour one defines a constant-value function, i.e. real-valued phase  $\phi_j$ ,  $j = 0 \dots \mathcal{N}$ . The general picture of NF spectrum structure of the finite-band solution is given in figure 3.3.

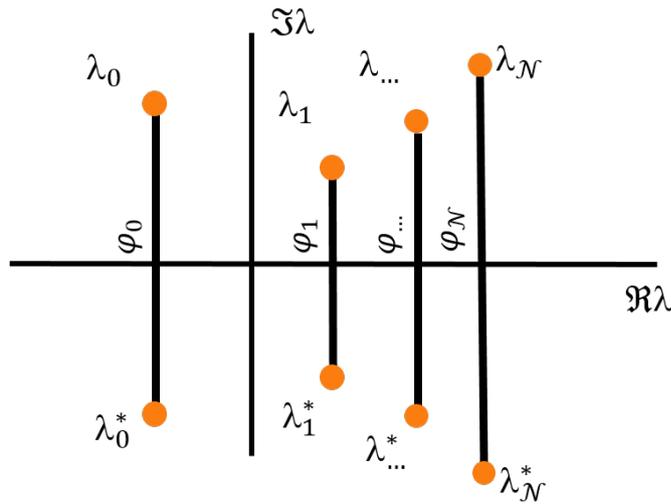


Figure 3.3: Visual representation of the NF image for the finite-band case. It consists of a main spectrum  $\lambda_j$ , and corresponding phases  $\phi_j$ , defined on the contours  $\Gamma_j = (\lambda_j, \lambda_j^*)$ .

Calculating the main spectrum points  $\lambda_j$  and the phases  $\phi_j$  from the given profile signal  $q(\tau)$  constitutes the direct part of the periodic NFT. According to the RHP formalism,

the goal here is to find its representation in terms of the solution of RHP. Here we have to restrict the problem to the case where  $q(\tau, \zeta)$  calculated via the solution of the inverse problem turns out to be periodic in  $\tau$  (for all  $\zeta$ ), with a certain period  $T$ .

The fundamental matrix made by these two solutions of ZSP system, eq. (3.5), being evaluated at one period  $T$  after the base point, is called the monodromy matrix:

$$\mathbf{M}(\lambda) = \Phi(\tau_0, \tau_0 + T, \lambda), \quad \Phi(\tau_0, \tau_0, \lambda) = I, \quad (3.45)$$

with  $\det \mathbf{M} = 1$  and the trace  $\Delta(\lambda) = \text{Tr} \mathbf{M}$ , called the discriminant, independent of the base point  $\tau_0$  because of the profile  $q(\tau)$  periodicity. In the Floquet theory for the Zakharov-Shabat operator with periodic coefficients, its Bloch solution  $\psi$  is determined as an eigenfunction of the operator of the shift by the period:

$$\psi(\tau + T; \lambda) = m(\lambda)\psi(\tau; \lambda), \quad (3.46)$$

where  $m(\lambda)$  is called the Floquet multiplier. Clearly,  $\psi$  is bounded on the whole line for those  $\lambda$ , for which  $|m(\lambda)| = 1$ . On the other hand, being a solution of the Zakharov-Shabat system,  $\psi$  is a combination of columns of  $\Phi$  at any time:

$$\begin{aligned} \psi(\tau_0 + T; \lambda) &= A\Phi_1(\tau_0, \tau_0 + T; \lambda) + B\Phi_2(\tau_0, \tau_0 + T; \lambda) \\ &= \mathbf{M} \begin{bmatrix} A \\ B \end{bmatrix}. \end{aligned}$$

Combining this with eq. (3.46), it follows that

$$m(\lambda) \begin{bmatrix} A \\ B \end{bmatrix} = \mathbf{M} \begin{bmatrix} A \\ B \end{bmatrix}, \quad (3.47)$$

i.e.,  $m(\lambda)$  is an eigenvalue of the monodromy matrix  $\mathbf{M}$ . Since  $\det \mathbf{M} = 1$ , two eigenvalues,  $m_+(\lambda)$  and  $m_-(\lambda)$ , are determined by  $\Delta(\lambda)$ :

$$m_{\pm}(\lambda) = \frac{\Delta(\lambda) \pm \sqrt{\Delta^2(\lambda) - 4}}{2}, \quad (3.48)$$

from which we can realise that  $\Delta(\lambda) \in [-2, 2]$  for the solution to be bounded. It follows that the continuous spectrum of the Zakharov-Shabat operator, which is where  $|m(\lambda)| = 1$ , can be characterized by the inequality  $\Delta^2(\lambda) \leq 4$ . For the end points  $\lambda = \lambda_j$ , we have  $\Delta(\lambda_j) = \pm 2$  and thus  $m(\lambda_j) = 1$  or  $m(\lambda_j) = -1$ , which, in view of (3.46), corresponds to the periodic or anti-periodic solutions of (3.5). Thus the end points of the spectral arcs (the main spectrum) can be found as the eigenvalues of the periodic or anti-periodic problem for (3.5) posed on the period interval.

Let  $\lambda$  belong to the spectrum associated to two Bloch solutions,  $\psi^+$  and  $\psi^-$ . To construct a periodic solution to the NLS, one can define squared eigenfunctions using these Bloch solutions as follows [131]:

$$\begin{aligned} f(\tau, \zeta; \lambda) &= -\frac{i}{2} (\psi_1^+ \psi_2^- + \psi_2^+ \psi_1^-), \\ g(\tau, \zeta; \lambda) &= \psi_1^+ \psi_1^-, \\ h(\tau, \zeta; \lambda) &= -\psi_2^+ \psi_2^-. \end{aligned}$$

Having a finite-band solution implies that these squared eigenfunctions are finite-order polynomials in  $\lambda$ . Straightforward calculations yield:

$$\begin{aligned} f^2 - gh &= -\frac{1}{4} W(\psi^+ \psi^-)^2 = \prod_{j=0}^{\mathcal{N}} (\lambda - \lambda_j) (\lambda - \lambda_j^*) = w^2(\lambda), \\ g(\tau, \zeta; \lambda) &= iq(t, \zeta) \times \prod_{j=1}^{\mathcal{N}} (\lambda - \mu_j(\tau, \zeta)), \end{aligned} \quad (3.49)$$

where  $W(\cdot)$  is the Wronskian and  $\mu_j(\tau, \zeta)$  are the auxiliary spectrum, which represents the evolution of the signal as it propagates through the fibre. To find  $\mu_j(\tau, \zeta)$  at any  $\zeta$  and  $\tau$ , one needs to solve a system of differential equations, see [131]. The auxiliary spectrum lies on a Riemann surface defined by eq. (3.44) [63, 131, 141].

In the RHP approach to the inverse periodic NFT, the determination of the corresponding phases  $\phi_j$  involves the off-diagonal elements of  $\mathbf{M}$ :  $M_{(12)}$  and  $M_{(21)}$ , evaluated at  $\Gamma_j$ , and the points of the auxiliary spectrum,  $\mu_j$  which are either double zeros of  $M_{(12)}$  or simple zeros of  $M_{(21)}$  that do not coincide with double zeros of the l.h.s. of (3.48). Namely, introducing the function:

$$P(\lambda) = \log \left( \sqrt{-\frac{M_{(12)}(\lambda)}{M_{(21)}(\lambda)} \prod_{j=1}^{\mathcal{N}} \frac{\lambda - \mu_j^*}{\lambda - \mu_j}} \right), \quad (3.50)$$

the phases  $\{\phi_j\}_{j=0}^{\mathcal{N}}$  are uniquely determined from the system of  $\mathcal{N} + 1$  linear equations ( $\phi = [\phi_0, \dots, \phi_{\mathcal{N}}]^T$ ):

$$\mathbf{K}\phi = b. \quad (3.51)$$

Here the entries of the  $(\mathcal{N} + 1) \times (\mathcal{N} + 1)$  matrix  $R(m, j) = \mathbf{R}_{(mj)}$  are defined by

$$\mathbf{R}_{(mj)} = \int_{\Gamma_j} \frac{x^m dx}{w^+(x)}, \quad j, m = 0, \dots, \mathcal{N}, \quad (3.52)$$

and the components  $b_m$  of the r.h.s. vector  $\mathbf{b}$  are as follows

$$b_m = -i \sum_{j=0}^{\mathcal{N}} \int_{\Gamma_j} \frac{P(x)x^m}{w^+(x)} dx, \quad m = 0, \dots, \mathcal{N}. \quad (3.53)$$

The algorithm (3.50)-(3.53) for determining  $\phi_j$  can be justified by using a series of deformations of RHP associated with NLS with periodic  $q$  in  $\tau$ . There is also a certain analogy with the case of the Korteweg-de Vries equation [100], which was used to derive the procedure above.

Ignoring the choice of the branch of the square root in the definition of  $P$ , eq. (3.50), the phases are reconstructed by modulo  $\pi$ , such that the phase's two values different by  $\pi$  cannot be distinguished.

### 3.2.2 Evolution of scattering data

The similarity between vanishing and periodic scattering data lies in the fact that they both consist of some static part and some dynamic part. In the vanishing boundary condition case, the static part is discrete eigenvalues, when the continuous spectrum and the spectral amplitudes change dynamically, see subsection 3.1.4. In the finite-band case, the static part is main spectrum points, and the auxiliary spectrum points, as well as the phases in RHP formalism, change dynamically.

The auxiliary spectrum, within Its and Kotlyarov formalism [59], satisfies the following partial differential equations:

$$\begin{aligned} \frac{\partial \mu_j}{\partial \tau} &= \frac{-2i\kappa_j w(\mu_j)}{\prod_{l \neq j} (\mu_j - \mu_l)} \\ \frac{\partial \mu_j}{\partial \zeta} &= -2 \left( \sum_{l \neq j}^{\mathcal{N}} \mu_l - \frac{1}{2} \sum_{j=0}^{\mathcal{N}} (\lambda_j + \lambda_j^*) \right) \frac{\partial \mu_j}{\partial \tau}, \end{aligned} \quad (3.54)$$

where  $\kappa_j \in \{-1, 1\}$  designates the particular sheet of the Riemann surface, defined by  $w(\lambda)$ , where the given auxiliary spectrum point lies on. Fortunately, one can use the

so-called Abel map transform, which reduces the localisation problem of the  $\mu_j$  in the Riemann surface to the localisation in the complex plane via the corresponding change of variables. However, as in this research I am not involving the auxiliary spectrum points as a part of the communication system, one can refer to the Abel mapping elsewhere [14].

The phases, appearing naturally as a part of the inverse problem, see subsection 3.3.5, are used in the expressions for the jump matrices,  $\hat{\mathbf{J}}_j(\tau, \zeta)$  having a linear dependence on  $t$  and  $z$  demonstrating explicitly the (quasi) periodic behaviour of  $\hat{\Phi}$  and thus  $q(\tau, \zeta)$ . The phases  $\phi_j$  change with the corresponding frequencies  $C^g$ , see eq. (3.86) as:

$$\phi_j(\zeta) = \phi_j(\zeta = 0) + (C_j^g - 2g_0)\zeta, \quad j = 0, \dots, \mathcal{N}. \quad (3.55)$$

### 3.2.3 Discussion

The finite-band approach introduces the characterisation of some class of NLS solutions with the finite number of components, known as main and auxiliary spectrum. In the focusing NLS case, studied here, the more direct approach replaces the auxiliary spectrum (with unclear localisation over the complex plane of the spectral parameter  $\lambda$  and nonlinear constrain atop) with a single real-values phase parameter. In this thesis, we provide an explicit way of determining phases from the periodic NLS solution. Apart from getting rid of the nonlinear coupling of the auxiliary spectrum, the introduction of phases provides an advantage of linear evolution with  $\zeta$  in NF domain, which allows employing these factors as data carriers.

The similarity between vanishing and periodic boundary condition cases is more direct for the forward NFT, when in both cases all definitions and properties are gathered around ZSP and its solution properties. Another expected similarity is between the solitonic eigenvalues in vanishing case, and main spectrum points in finite-band one, which is not so transparent. There is no clear correspondence between the main spectrum point-phase pair values and the structure of the corresponding mode in the time domain.

## 3.3 Inverse NFT

### 3.3.1 General remarks

In this section I describe various method for INFT: the procedure which maps the given NF picture (for certain boundary conditions) to the profile in  $(\tau, \zeta)$  domain, i.e. unique solution of NLS. For the vanishing boundary there more comprehensive studies have been done, see review [135], in particular, because this case is simpler and has application beyond optical communication. For example, IST technique is widely used for computation of fibre Bragg grating profiles, see [66, 38].

One of the great achievements of last years is the development of fast performing numerical algorithms for INFT for both vanishing [140] and periodic cases [141], with the proclaimed complexity of  $O(N_q \cdot \log^2 N_q)$  for  $N_q$  signal samples in time domain. However, the numerical complexity of particularly inverse transformation on the cycle of back-to-back mapping is still a severe bottleneck in developing efficient transmission NFT-based communication schemes.

In the vanishing boundary condition case, the chosen INFT method depends on the presence or absence of the specific NF spectrum components, i.e. continuous or discrete spectra. In periodic boundary condition case, the chosen method depends on a way of the definition and association of the auxiliary component.

In communications application, the main requirements to the numerical method is not usually a mathematical purity of the derivations and expression, or elegant representation in the closed-form expressions, but the computational complexity of the numerical methods

used on the Tx and the Rx. Therefore, the most promising method, which I used in this research for all simulation, is the so-called inverse layer peeling. It literally performs backwards the operation for matrix multiplications, done in the transfer-matrix approach, see 4.2. The main reason of using it is that its fast version, developed by Wahls and Vaibhav [142], has the complexity  $O(N_q \log^2 N_q)$ , like the fast direct NFT algorithm [140].

Here I have to do some remarks about the motivation and generality for the described method for the finite-band NLS solution. The periodicity issue in the integrable systems has been the subject of study for a long time [17, 58, 76, 96, 152], and their vast application area ranging from water gravity waves to optical turbulence, has encouraged researchers to develop a solid mathematical base for their description [5, 21, 110]. Specifically for the periodic case of NLS solutions, the various studies have been done [14, 21, 59, 63, 99].

There are three principal approaches to parametrising and generation a periodic solution of NLS by means of IST. They introduce a different kind of spectral data, so the way of performing of INFT should be aligned with the procedure of direct NFT. The first approach, described by Its and Kotlyarov in [59, 73], and its alternative version by Ma and Ablowitz [96, 97], involves the notions of main and auxiliary spectra in their original meaning. While the main spectrum points are more general notion, being naturally involved in RHP approach as well, see figure 3.3, the auxiliary spectrum can be defined in a various way, with the corresponding changes in the equations which describe their evolution. For Ma and Ablowitz method, the signal reconstruction is simpler, however, the evolution of the auxiliary spectra is more clear for Its and Kotlyarov definitions, so the full cycle of transforms has roughly similar complexity [141].

The main drawback of this method that the auxiliary spectra, chosen there, evolves according to the coupled partial differential equations, so it does not provide the effective linearisation of NLS channel, what is the very advantage of NFT. Also, not every choice of the scattering data would lead to the eligible solution of NLS, and for Ma and Ablowitz definition it is not possible to write the closer-form conditions for auxiliary spectral values. However, some research has been done in this direction, for example, the developing of effective numerical algorithms [141].

Another method is the so-called algebro-geometric approach, where a periodic NLS solution is expressed in terms of meromorphic functions, differentials and integrals on hyperelliptic Riemann surfaces of the finite genus, see [14]. The Riemann surface is defined by the structure of the main spectrum, see eq. (3.44). Then one can introduce the canonical basis of oriented circles  $(a_j, b_j)$ , and define the associated period matrix  $B$  and the Abel integrals, fixed by their particular behaviour at  $\infty^\pm$ . Basing on the period matrix  $B$ , one can explicitly find the desired NLS solution via the computation of the so-called Riemann theta functions, which effectively are multidimensional Fourier series with, in general, random mode frequencies. However, having, in theory, the mathematically explicit expression for  $q(\tau, \zeta)$ , in numerics, the noticeable numerical error is gained by truncating of the grid in the frequency-wavenumber domain, which can be controlled to some extent by considering the contribution of different terms along with the series [29]. In addition, the numerical complexity grows exponentially with the number of bands  $\mathcal{N}$ , which is determining, in turn, the potential capacity of the communication system. Considering the so-called the hyperfast method of evaluating the Riemann theta function [110], the complexity of constructing a periodic signal with  $N_q$  time samples over the  $M$  samples in NF domain can be estimated as:

$$O\left(N_q \mathcal{N} (2M + 1)^\mathcal{N} \log(2M + 1)\right). \quad (3.56)$$

Although the latter property helps decrease the computational complexity, it remains considerably heavy, especially when compared to its alternative that is rendered by the RHP approach. For the full mathematical description of the algebro-geometric method,

the reader can refer to [63, 72].

Here we must comment on the issue of applicability of the finite-band solution in the optical communication application. The finite-band solutions are well-known for the existence of modulational instability [65, 131]. Recalling that waveforms represent the optical signal, it is highly undesirable, because developing of high-amplitude (power) profiles can be catastrophic for the physical components of the system, i.e. fibres, amplifiers and transceivers. The development of modulational instability is known, for example, as a responsible effect for rogue waves in water.

Because we deal with the artificially constructed structure of bands (as the information modulation is performed in NF domain), we can control the achievable signal power on the Tx. Then, we always have to remember that the practical channel imposes deviations from the pure integrable NLS, and in particular, it includes the attenuation (see subsection 2.1.4). These imperfections of the integrability model suppress the development of high amplitude waves. Also, the propagation distances considered in this research (being in the domain of long-haul communications), provide not large values in normalised (solitonic) units of variable  $\zeta$ . Therefore, we rather deal in the regime of  $\zeta \lesssim 1$  then  $\zeta \ll 1$ , so even if the signal contains some potentially rising modes, they would not develop to noticeable amplitudes. These three reasons are sufficient to not consider possible modulational instability as an impactful factor.

Because of the listed drawbacks and problems, in the application for the purposes of the communications, this research followed the third possible approach in IST for periodic boundary conditions, i.e. RHP approach accompanied with the additional constrains over the main spectrum which leads to the exactly periodic solution. In this thesis, I do not use these alternative approaches to the application for optical communication, therefore the full description of the mathematical procedures behind them the reader can find elsewhere [141].

### 3.3.2 Gelfand-Levitan-Marchenko equations

In the vanishing boundary condition case, the INFT is defined as an operation which maps the combination of the scattering data in NF domain to temporal domain, i.e. generates the corresponding solution of NLS. The most comprehensive method, which naturally involves both continuous and discrete data, is to solve the Gelfand-Levitan-Marchenko equation (GLM) [2, 37, 78].

This method indirectly takes the scattering data as an equation term, defining the integral equation *kernel* via left reflection coefficient  $l(\lambda)$  from eq. (3.18), and left spectral amplitudes  $d_j$  from eq. (3.21):

$$\Omega_L(\theta) = \int_{-\infty}^{\infty} \frac{d\lambda}{2\pi} e^{-i\lambda\theta} l(\lambda) + (-i) \sum_j d_j e^{-i\lambda_j\theta} \quad (3.57)$$

so adding the discrete spectrum modes to explicitly IFT of the continuous spectrum (with certain normalisation), mapping NF space of parameter  $\lambda$  to its Fourier counterpart, parametrized by variable  $\theta$ . GLM is given as a system of coupled integral equations for unknown functions  $A_{1,2}(s, s')$ :

$$\begin{aligned} A_1^*(x, y) + \int_{-\infty}^x dz A_2(x, z) \Omega_L(y + z) &= 0 \\ -A_2^*(x, y) + \Omega_L(x + y) + \int_{-\infty}^x dz A_1(x, z) \Omega_L(y + z) &= 0, \end{aligned} \quad (3.58)$$

for  $-x \leq y \leq x$  and  $0 \leq x$ . Note that GLM effectively finds the solution on the positive half-line and it should be treated correspondingly. The solutions of NLS, which is unambiguously determined by the function  $\Omega_L(\theta)$ , is given as  $q(\tau) = -2A_2(\tau, \tau - 0)$ . The

derivations in [78] and [38] use different integration limits because ref. [38] specifically addresses the generation of fibre Bragg grating profiles. Also note that in practice one has to properly pick the spectral amplitudes and eigenvalues to locate the solitonic modes within the dedicated time window.

The inverse mapping between the input of GLM  $\Omega_L$  and left reflection coefficient  $l(\lambda)$  (for empty discrete spectrum set) is given as:

$$l(\lambda) = \int_0^\infty d\theta \Omega_L(\theta) e^{i\lambda\theta}. \quad (3.59)$$

The idea of using GLM as a tool for INFT possesses a couple of disadvantages which I want to list here. First, it involves additional IFT mapping of scattering data, which can be undesirable and may introduce a numerical error even for smooth profiles. For example, the famous issue of the Gibbs phenomena, known for rectangular profiles, being the case for the NFT-based transmission, because the Nyquist modulation is often used there [32, 80]. Together with this, the appearance of the solitonic modes creates divergence of the entries  $\sim e^{-i\lambda_j\theta}$ , which dramatically increase for long potentials and/or high-energy solitonic components. In addition, the numerical methods used for the computation of the profiles accumulate the numerical error, since each iteration involves the previously generated samples [38].

### 3.3.3 Darboux transform

The Darboux transform is a method applicable for various integrable equation in the context of the existence of coherent structures in their solutions [49, 123]. In particular, it allows embedding of solitonic components to the existent solution of NLS via correct redefinition of the Jost functions  $\psi$  from eq. (3.5). In particular, it is known that for  $N_s$ -soliton profile (with no dispersive modes) is described by the following NF spectrum [144]:

$$b(\lambda) = 0, \quad \forall \lambda \in \mathbb{R}; \quad a(\lambda) = \prod_{j=1}^{N_s} \frac{\lambda - \lambda_j}{\lambda - \lambda_j^*}, \quad (3.60)$$

however, in practice, one has to properly pick the spectral amplitudes components in order to locate the solitons inside computational time window of interest.

In this research, I mainly used Darboux transform when generating signals with certain bound states, because it is rather simple and transparent, and it does not involve the additional artificial transform to Fourier image of NF space, i.e. function  $\Omega_l$  from eq. (3.57). The Darboux transform uses the properties of AKNS system, eqs. (3.6-3.7). Let us define a matrix of Jost functions, which the first column satisfies ZSP for the seed potential  $q(\tau)$  as:

$$\frac{\partial}{\partial \tau} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \hat{\mathbf{T}} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \quad \Psi = \begin{pmatrix} \psi_1 & \psi_2^* \\ \psi_2 & -\psi_1^* \end{pmatrix}, \quad \Psi = \Psi(\tau, \lambda), \quad (3.61)$$

and a complex  $\lambda_0$ ,  $\Im \lambda > 0$ , one wants to be a discrete eigenvalue of the updated potential  $\tilde{q}(\tau)$ . Rest of the scattering data of the updated potential  $\tilde{q}(\tau)$  should coincide with the scattering data of the seed one  $q(\tau)$ . Let us update the Jost functions matrix  $\Psi$  as  $\Sigma = \Psi \Lambda_0 \Psi$ , where

$$\Lambda_0 = \begin{pmatrix} \lambda_0 & 0 \\ 0 & \lambda_0^* \end{pmatrix}. \quad (3.62)$$

Define the updated signal as:

$$\tilde{q}(\tau) = q(\tau) - 2i(\lambda_0 - \lambda_0^*) \frac{\psi_2^*(\tau, \lambda_0) \psi_1(\tau, \lambda_0)}{|\psi_1(\tau, \lambda_j)| + |\psi_2(\tau, \lambda_j)|}, \quad (3.63)$$

so its scattering data will modify as follows:

$$\begin{aligned} \tilde{a}(\lambda) &= a(\lambda) \frac{\lambda - \lambda_0}{\lambda - \lambda_0^*}, & \tilde{b}(\lambda) &= b(\lambda), & \forall \lambda \in \mathbb{R} \\ \tilde{c}(\lambda_{j \neq 0}) &= c(\lambda_{j \neq 0}) \frac{\lambda_{j \neq 0} - \lambda_0^*}{\lambda_{j \neq 0} - \lambda_0}. \end{aligned} \quad (3.64)$$

For the formal proof, the reader can refer to [10]. Note that the embedding eigenvalue  $\lambda_0$  must not be a zero of the scattering function  $a(\lambda)$  before the application of Darboux transform, otherwise the procedure will remove the eigenvalue from the existent set of the discrete eigenvalues.

This method possesses several advantages that make it a promising and universal for discrete eigenvalues modulation in communications. It takes all eigenvalues in turn, so they can be iteratively embedded one by one. It means that the numerical complexity of the algorithm scales linearly with the number of the bound states. However, it affects the expressions for the scattering function  $a(\lambda)$ , which can be critical for applications.

### 3.3.4 RHP for vanishing boundary condition

The framework of RHP is more universal than is used by the NFT community. In particular, it allows defining a certain RHP for IST starting with the scattering data for the vanishing case. In this subsection, I provide the procedure of building a vanishing solution of NLS via the solution of the corresponding RHP. Note here that it corresponds to the case without solitonic degrees of freedom, i.e. with empty discrete spectrum. The results in this subsection form a preliminary research in using RHP approach for vanishing case, which can be convenient for numerics, fast and reliable. For test purposes, I used sech (Satsuma-Yajima) potential [126, 133] with explicit expressions for the scattering data:

$$\begin{aligned} q(\tau) &= -iA \operatorname{sech} \tau, \\ a(\lambda) &= \frac{\Gamma^2(1/2 - i\lambda)}{\Gamma(1/2 - A - i\lambda)\Gamma(1/2 + A - i\lambda)}, & b(\lambda) &= \frac{i\pi}{A \operatorname{cosh}(\pi\lambda)\Gamma(A)\Gamma(-A)}, \end{aligned} \quad (3.65)$$

having in mind that the potential  $L_1$ -norm (controlled by amplitude  $A$ ) should be sufficiently small (less than  $\pi/2$ ) to not rise the solitonic modes. For the numerical solution of RHP here and in the following research, I extensively used Wolfram Mathematica RH-Package by Olver [107].

In general, RHP is a boundary value problem in the complex plane, typically consisting of finding a unique (matrix-valued) function  $\mathbf{M}$  from the given (matrix-valued) function  $\mathbf{J}$  (having also in mind its domain, the (oriented) contour  $\Gamma$ , where it is defined), which satisfies the following conditions:

- (i) out of contour analyticity: the solution of RHP  $\mathbf{M}$  is analytic for  $\lambda \in \mathbb{C} \setminus \Gamma$ .
- (ii) asymptote at infinity:  $\mathbf{M}(\lambda) \rightarrow I$  as  $\lambda \rightarrow \infty$ .
- (iii) the jump condition when approaching the contour  $\Gamma$  from either side:

$$\mathbf{M}^+(\lambda) = \mathbf{M}^-(\lambda)\mathbf{J}(\lambda). \quad (3.66)$$

The problem can be easily parametrised, as soon as the conditions (especially asymptotic behaviour) are ensured for all values of parameters. For 1+1 NLS, eq. (2.15), one can define RHP as

$$\mathbf{M}^+(\lambda, \tau, \zeta) = \mathbf{M}^-(\lambda, \tau, \zeta)\mathbf{J}(\lambda, \tau, \zeta). \quad (3.67)$$

The corresponding NLS solution can be found via the asymptotic behaviour as follows:

$$\begin{aligned}\mathbf{M}(\lambda, \tau, \zeta) &= I + \lambda^{-1} \mathbf{M}^{(i)}(\tau, \zeta) + O(\lambda^{-2}), \quad \lambda \rightarrow \infty, \\ q(\tau, \zeta) &= 2i \mathbf{M}^{(i)}_{1,2}(\tau, \zeta).\end{aligned}\quad (3.68)$$

In the vanishing boundary condition case, for  $\zeta = 0$ , one can start with either continuous spectrum function  $r(\lambda)$ , or  $b(\lambda)$ , as they both are sufficient to represent the full set of spectral data. In first case, the jump matrix is given by:

$$\mathbf{J}(\lambda, \tau) = \exp(-i\lambda\tau\sigma_3) \begin{pmatrix} 1 & r^*(\lambda^*) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ r(\lambda) & 1 \end{pmatrix} \exp(i\lambda\tau\sigma_3), \quad \lambda \in \mathbb{R}, \quad (3.69)$$

when for  $b(\lambda)$  we get simpler expression

$$\mathbf{J}(\lambda, \tau) = \exp(-i\lambda\tau\sigma_3) \begin{pmatrix} 1 & b(\lambda) \\ b^*(\lambda^*) & 1 \end{pmatrix} \exp(i\lambda\tau\sigma_3), \quad \lambda \in \mathbb{R}, \quad (3.70)$$

however, because of the analyticity issue, these formulations may have problems, if we deform RHP to step in the upper half-plane.

One of the advantage of RHP approach is a possibility to use different RHP to get the same NLS solution. In particular, the jump matrix of the abovestated RHP decays not sufficiently rapidly (as  $1/\lambda$ ) at  $\lambda \rightarrow \pm\infty$ . It is not convenient for numerical solutions, in particular, because the contour  $\Gamma$  is artificially truncated on some finite extent and is sampled with finite resolution. The alternative way is the following: one can split the jump matrix (3.69) to get the decaying behaviour due to exponential wrapping by infinitesimal shift  $\epsilon$  to upper and lower half-planes. The resulting jump matrices and their contours are given as:

$$\begin{aligned}\mathbf{J}_1(\lambda, \tau) &= \begin{pmatrix} 1 & 0 \\ e^{2i\lambda\tau} r(\lambda) & 1 \end{pmatrix}, \quad \lambda \in (-\infty + i\epsilon, \infty + i\epsilon), \\ \mathbf{J}_2(\lambda, \tau) &= \begin{pmatrix} 1 & e^{-2i\lambda\tau} r^*(\lambda^*) \\ 0 & 1 \end{pmatrix}, \quad \lambda \in (-\infty - i\epsilon, \infty - i\epsilon).\end{aligned}\quad (3.71)$$

The reconstruction of the test potential evidences the advantage of the jump matrix splitting, as in eq. (3.71), returning more accurate results, see figure 3.4. However, usage of the jump matrix defined by  $b(\lambda)$ , eq. (3.70), performs better than any of jump matrices defined via  $r(\lambda)$ . This can be promising method for the becoming popular  $b$ -modulation approach, since it does not require to evaluate function  $a(\lambda)$  from  $b(\lambda)$  on the inverse NFT stage.

For  $\zeta > 0$ , the above RHP is inconvenient for numerical implementation, in particular, due to consideration of the stable points of the exponential wrapping. The not-deformed RHP is defined by the jump matrix similar to eq. (3.69):

$$\mathbf{J}(\lambda, \tau) = \exp(-i\lambda\tau\sigma_3) \begin{pmatrix} 1 & r^*(\lambda^*) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ r(\lambda) & 1 \end{pmatrix} \exp(i\lambda\tau\sigma_3), \quad \lambda \in \mathbb{R}, \quad (3.72)$$

so the (real) stationary point is  $\lambda_s = -\tau/4\zeta$ . Then the initial contour is split on diagonal sections to get the curves where the jump matrices sufficiently fast converge to identity matrix getting the topology as in figure 3.5.

The modification of the contours from fig. 3.5(b), provides the jump matrices given as:

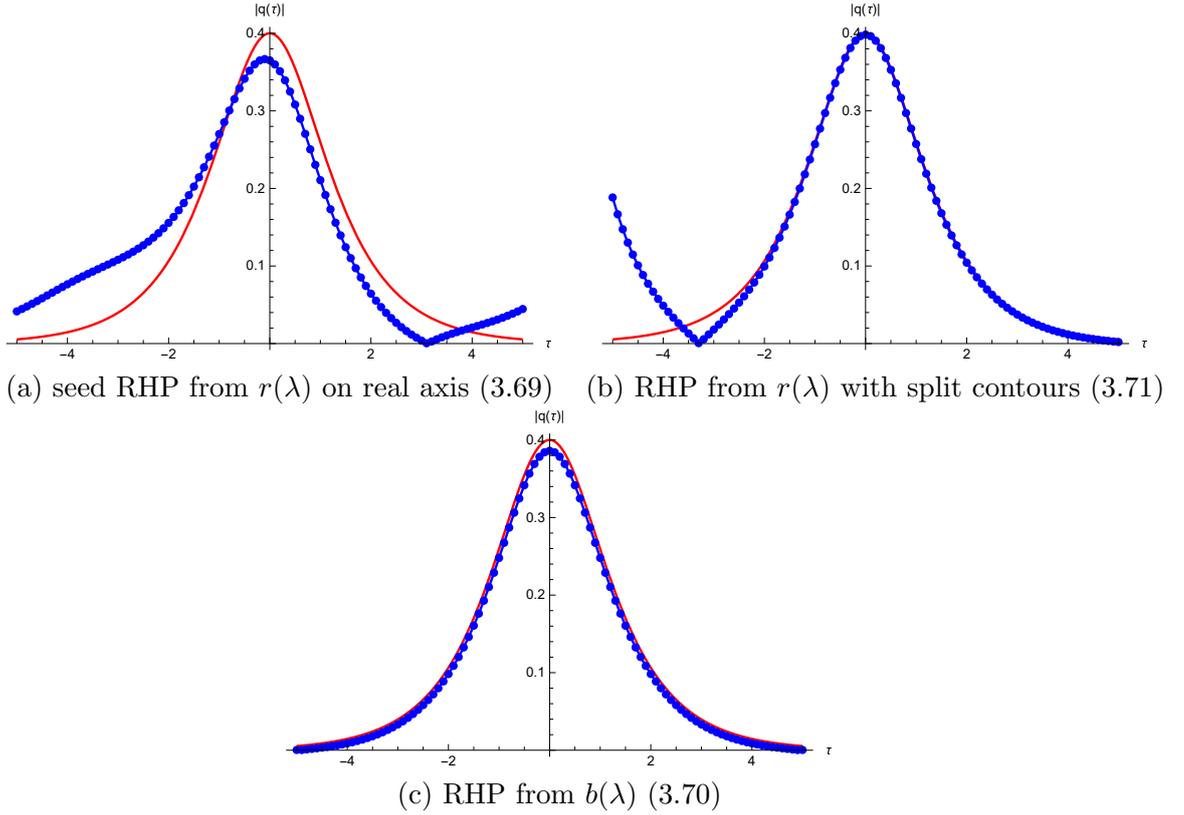


Figure 3.4: The sech potential reconstruction via RHP for vanishing boundary condition via different jump matrices expressions. The jump matrices computational domain is  $\Gamma = (-10, 10)$  with 100 sampling points.

$$\begin{aligned}
 \mathbf{H}_1(\lambda) &= \begin{pmatrix} 1 & 0 \\ r(\lambda) & 1 \end{pmatrix}, \\
 \mathbf{H}_2(\lambda) &= \begin{pmatrix} 1 & \frac{r^*(\lambda^*)}{r(\lambda)r^*(\lambda^*)+1} \\ 0 & 1 \end{pmatrix}, \\
 \mathbf{H}_3(\lambda) &= \begin{pmatrix} 1 & 0 \\ -\frac{r(\lambda)}{r(\lambda)r^*(\lambda^*)+1} & 1 \end{pmatrix}, \\
 \mathbf{H}_4(\lambda) &= \begin{pmatrix} 1 & -r^*(\lambda^*) \\ 0 & 1 \end{pmatrix}, \\
 \mathbf{H}_5(\lambda) &= \begin{pmatrix} r(\lambda)r^*(\lambda^*)+1 & 0 \\ 0 & \frac{1}{r(\lambda)r^*(\lambda^*)+1} \end{pmatrix},
 \end{aligned} \tag{3.73}$$

and

$$\mathbf{J}_m(\lambda, \tau, \zeta) = \exp[-(i\lambda\tau + 2i\lambda^2\zeta)\sigma_3]\mathbf{H}_m(\lambda) \exp[(i\lambda\tau + 2i\lambda^2\zeta)\sigma_3]. \tag{3.74}$$

In the deformation from fig. 3.5(c) we additionally avoid the stationary point to suppress possible singularities. To use RHP in this form, one should first define:

$$\delta(\lambda) = \exp\left(\frac{1}{2\pi i} \int_{-\infty}^{\lambda_s} \frac{\log(r(s)r^*(s^*)+1)}{s-\lambda} ds\right), \quad \mathbf{\Delta}(\lambda) = \begin{pmatrix} \delta(\lambda) & 0 \\ 0 & \frac{1}{\delta(\lambda)} \end{pmatrix}. \tag{3.75}$$

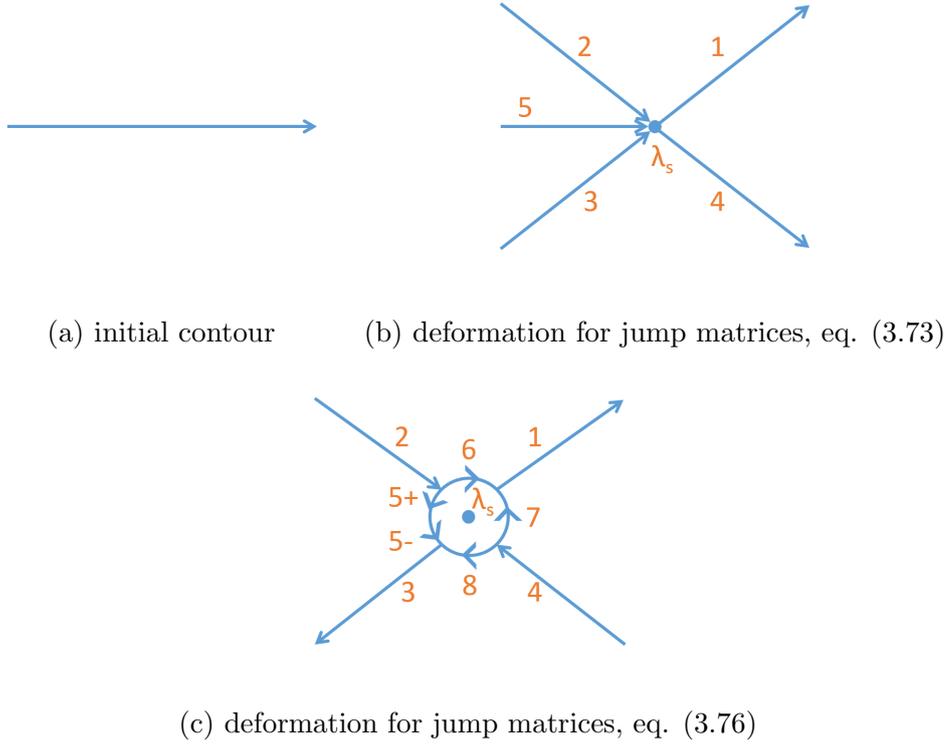


Figure 3.5: The contours  $\Gamma$ , used for the RHP in the vanishing boundary condition case for nonzero  $\zeta$  to deal with jump matrices convergence.

The corresponding jump matrices are given as:

$$\begin{aligned}
 \mathbf{H}_1(\lambda) &= \mathbf{\Delta}(\lambda) \begin{pmatrix} 1 & 0 \\ r(\lambda) & 1 \end{pmatrix} \mathbf{\Delta}^{-1}(\lambda), \\
 \mathbf{H}_2(\lambda) &= \mathbf{\Delta}(\lambda) \begin{pmatrix} 1 & \frac{r^*(\lambda^*)}{r(\lambda)r^*(\lambda^*)+1} \\ 0 & 1 \end{pmatrix} \mathbf{\Delta}^{-1}(\lambda), \\
 \mathbf{H}_3(\lambda) &= \mathbf{\Delta}(\lambda) \begin{pmatrix} 1 & 0 \\ -\frac{r(\lambda)}{r(\lambda)r^*(\lambda^*)+1} & 1 \end{pmatrix} \mathbf{\Delta}^{-1}(\lambda), \\
 \mathbf{H}_4(\lambda) &= \mathbf{\Delta}(\lambda) \begin{pmatrix} 1 & -r^*(\lambda^*) \\ 0 & 1 \end{pmatrix} \mathbf{\Delta}^{-1}(\lambda), \\
 \mathbf{H}_{5+}(\lambda) &= \mathbf{\Delta}(\lambda) \begin{pmatrix} 1 & \frac{r^*(\lambda^*)}{r(\lambda)r^*(\lambda^*)+1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -r(\lambda) & 1 \end{pmatrix}, \\
 \mathbf{H}_{5-}(\lambda) &= \mathbf{\Delta}(\lambda) \begin{pmatrix} 1 & 0 \\ -\frac{r(\lambda)}{r(\lambda)r^*(\lambda^*)+1} & 1 \end{pmatrix} \begin{pmatrix} 1 & r^*(\lambda^*) \\ 0 & 1 \end{pmatrix}, \\
 \mathbf{H}_6(\lambda) &= \begin{pmatrix} 1 & 0 \\ r(\lambda) & 1 \end{pmatrix} \mathbf{\Delta}^{-1}(\lambda), \\
 \mathbf{H}_7(\lambda) &= \mathbf{\Delta}(\lambda), \\
 \mathbf{H}_8(\lambda) &= \begin{pmatrix} 1 & -r^*(\lambda^*) \\ 0 & 1 \end{pmatrix} \mathbf{\Delta}^{-1}(\lambda),
 \end{aligned} \tag{3.76}$$

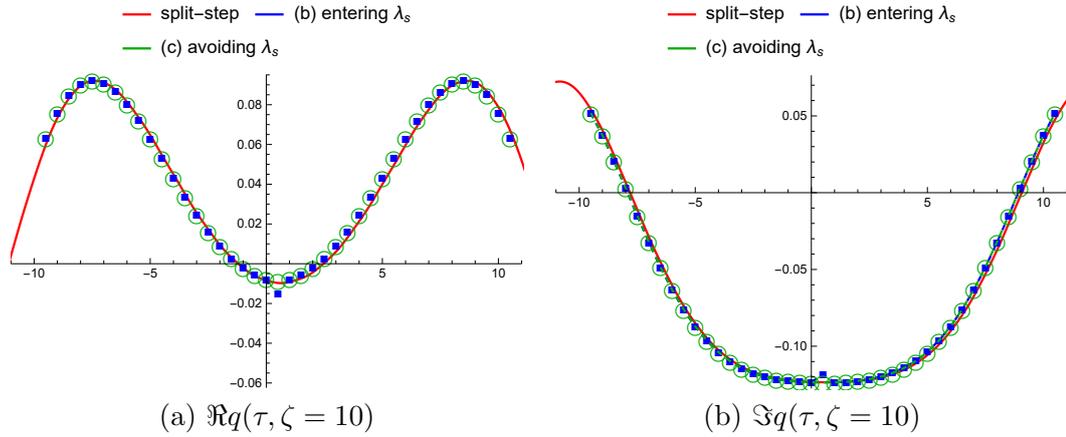
and

$$\mathbf{J}_m(\lambda, \tau, \zeta) = \exp[-(i\lambda\tau + 2i\lambda^2\zeta)\sigma_3]\mathbf{H}_m(\lambda)\exp[(i\lambda\tau + 2i\lambda^2\zeta)\sigma_3]. \quad (3.77)$$

The exponential enveloping provides the exponential decay of on the contours 1-4, so they can be truncated at some close point.

The resulting waveforms and the comparison with the same initial signal split-step propagation are provided in fig. 3.6. In the flexibility of the RHP approach it is equivalently simple to find the solution of NLS for any  $\zeta$ . The computation demonstrates that the point  $q(\tau = 0)$  escapes from the proper location, when the stationary point is not wrapped.

Figure 3.6: The numerical solution of RHP in the vanishing boundary condition case for  $\zeta = 10$  for two contours deformations.



I provide these expressions for illustrational purpose, to demonstrate the application of RHP concept to vanishing condition case. During my research in this area, I did not find a proper way how to determine the jump matrices by given analytic formula for  $r(\lambda)$ , eq. 3.65, but by random modulated waveform.

### 3.3.5 RHP in finite-band case

The main result of this thesis consists of the developing of the comprehensive back-to-back systems based on the performing IST by means of the solution of RHP. The main challenge here, apart of bulky mathematical formulation and not-optimised numerical methods, is the requirement to ensure the finite-band topology in NF domain, which will provide an exactly periodic profile in the temporal domain. The mathematical foundation for this work is done in collaboration with Prof Shepelsky and provided in two journal publications [J1, SJ1] and several conference contributions [C5, C6, C7].

In this subsection, I describe a procedure of inverse NFT from the given NF spectrum, in this formalism fully defined by  $\mathcal{N} + 1$  bands, i.e. main spectrum, and  $\mathcal{N} + 1$  real values phases, chosen from the range  $[0, 2\pi]$ . The corresponding direct transform is described in subsection 3.2.1.

Going one step forward, we will construct a solution to eqs. (3.1-3.2) and use it to make up  $q(\tau, \zeta)$  periodic in  $\tau$ . For doing this, we parametrise a family of solutions of eqs. (3.1-3.2) with a set of quantities that change with  $z$  in known manner.

A distinguishing feature of the RHP that one can adopt here is that its jump matrices are constants (w.r.t.  $\lambda$ ) on each connected part of the jump contours. The jump contour, in our particular case, connects the points of the so-called main spectrum [63],  $\{\lambda_j\}_{j=0}^{\mathcal{N}}$  with  $\lambda_j \in \mathbb{C} \setminus \mathbb{R}$ , and their conjugate, for more explicit details. Each individual arc number  $i$ , belonging to our contour, connects  $\lambda_i$  and its conjugate.

Next we have to define the jump matrix  $\mathbf{J}_j$  associated with each arc, by defining the values of real phases,  $\{\phi_j\}_{j=0}^{\mathcal{N}}$ , with  $\phi_j \in (0, 2\pi)$  (an analogue of vector  $\mathbf{D}$  in the algebro-geometric approach), see eq. (3.79) below. Together with these real-valued phases, we chose the poles of  $\mathfrak{Y}$ , which are the last required parameters to completely construct it, more details to follow. By solving the RHP, we find the so-called (planar) Baker-Akhiezer function  $\Phi(t, z, \lambda)$ , which is a special solution of both (3.1-3.2) defined on the complex  $\lambda$ -plane. So by means of RHP we associate the Baker-Akhiezer function with picture in  $\lambda$ -plane, i.e. with the given set of  $\mathcal{N} + 1$  complex and  $\mathcal{N} + 1$  real parameters,  $\{\{\lambda_j\}_{j=0}^{\mathcal{N}}, \{\phi_j\}_{j=0}^{\mathcal{N}}\}$  [72]. This Baker-Akhiezer function is uniquely characterised by the following conditions.

- (i) For any  $\tau$  and  $\zeta$ ,  $\Phi(\tau, \zeta, \lambda)$  is analytic,  $2 \times 2$ -valued function of  $\lambda \in \mathbb{C} \setminus \Gamma$ ,  $\Gamma$  being the union of arcs,  $\Gamma = \cup_{j=0}^{\mathcal{N}} \Gamma_j$ , where  $\Gamma_j = (\lambda_j, \lambda_j^*)$  is the particular arc connecting  $\lambda_j$  and its complex conjugate.
- (ii) The limiting values  $\Phi^\pm$  of  $\Phi$ , as  $\lambda$  approaches the both sides of  $\Gamma_j$ , are related through the jump conditions:

$$\Phi^-(\tau, \zeta, \lambda) = \Phi^+(\tau, \zeta, \lambda) \mathbf{J}_j, \quad \lambda \in \Gamma_j, \quad j = 0, \dots, \mathcal{N}, \quad (3.78)$$

$$\mathbf{J}_j = \begin{bmatrix} 0 & ie^{-i\phi_j} \\ ie^{i\phi_j} & 0 \end{bmatrix}. \quad (3.79)$$

Note that, as we said before, the jump matrix number  $j$  contains the phase  $\phi_j$  of  $j$ -th partial nonlinear mode. These are exactly the parameters that we will further use for the modulation.

- (iii) As  $\lambda \rightarrow \infty$ , the limiting value of Baker-Akhiezer function satisfies:

$$\Phi(\tau, \zeta, \lambda) = \left[ \mathbf{I} + O(\lambda^{-1}) \right] e^{-i(\lambda\tau + 2\lambda^2\zeta)\sigma_3}. \quad (3.80)$$

The statements above define our RHP written here for the matrix function  $\Phi(\tau, \zeta, \lambda)$ . Having solved the respective RHP, the associated genus- $\mathcal{N}$  solution  $q(\tau, \zeta)$  of NLS is given by the simple expression:

$$q(\tau, \zeta) = 2i \Phi_{1,2}^{(i)}(\tau, \zeta), \quad (3.81)$$

where the subscript  $\{1, 2\}$  stands for the corresponding matrix entry, and the quantity  $\Phi^{(i)}(\tau, \zeta)$  is defined through the limiting relation involving the Baker-Akhiezer function:

$$\Phi_{1,2}^{(i)}(\tau, \zeta) = \lim_{\lambda \rightarrow \infty} \lambda \left( \Phi(\tau, \zeta, \lambda) e^{i(\lambda\tau + 2\lambda^2\zeta)\sigma_3} - \mathbf{I} \right). \quad (3.82)$$

Thus, to construct the time domain waveform given the set of parameters  $\{\{\lambda_j\}_{j=0}^{\mathcal{N}}, \{\phi_j\}_{j=0}^{\mathcal{N}}\}$ , we need to define the jump matrices  $\mathbf{J}_j$  via eq. (3.79), and solve RHP (3.78)–(3.80) to arrive at  $\Phi(t, z, \lambda)$  and, eventually, at the sought  $q(\tau, \zeta)$  profile. In turn, the initial RHP (3.78)–(3.80) can be transformed to that of the form having the standard normalisation condition as  $\lambda \rightarrow \infty$  and the jump matrices  $\hat{\mathbf{J}}_j(t, z)$  independent of  $\lambda$  on each arc of the contour.

The finite-genus solutions are quasi-periodic in  $\tau$  but not, in general, periodic. In order to arrive at periodic NLS solutions, the spectral data  $\{\lambda_j\}_{j=0}^{\mathcal{N}}$  have to satisfy a system of transcendental equations ensuring that all the frequencies  $C_j^f$ ,  $j = 1, \dots, \mathcal{N}$  supplemented by  $f_0$ , are commensurable [14]. This leads to the reformulation of the original RHP, where these frequencies appear naturally [72]. One may change the variables as:

$$\Phi(\tau, \zeta, \lambda) = e^{(if_0\tau + ig_0\zeta)\sigma_3} \mathbf{M}(\tau, \zeta, \lambda) e^{-i(f(\lambda)\tau + ig(\lambda)\zeta)\sigma_3}, \quad (3.83)$$

where:

1.  $f(\lambda)$  and  $g(\lambda)$  are scalar functions analytic in  $\mathbb{C} \setminus \Gamma$  satisfying the following conditions:
  - (i)  $f(\lambda) = \lambda + f_0 + O(1/\lambda)$  and  $g(\lambda) = 2\lambda^2 + g_0 + O(1/\lambda)$  as  $\lambda \rightarrow \infty$ , with some constants  $f_0$  and  $g_0$ ; (ii) the limiting values of  $f$  and  $g$  across  $\Gamma$  are related by

$$f_+(\lambda) + f_-(\lambda) = C_j^f, \quad g_+(\lambda) + g_-(\lambda) = C_j^g, \quad j = 0, \dots, \mathcal{N}, \quad (3.84)$$

with some real constants  $C_j^f$  and  $C_j^g$  (here  $C_0^f = C_0^g = 0$ ).

2. The matrix  $\mathbf{M}$  is the solution of RHP with
  - (i) the jump conditions

$$\mathbf{M}^-(\tau, \zeta, \lambda) = \mathbf{M}^+(\tau, \zeta, \lambda) \mathbf{G}_j(\tau, \zeta), \quad \lambda \in \Gamma_j, \quad (3.85)$$

where

$$\mathbf{G}_j(\tau, \zeta) = \begin{pmatrix} 0 & ie^{-i(C_j^f \tau + C_j^g \zeta + \phi_j)} \\ ie^{i(C_j^f \tau + C_j^g \zeta + \phi_j)} & 0 \end{pmatrix}, \quad (3.86)$$

and

- (ii) the normalization condition  $\mathbf{M} \rightarrow I$  as  $\lambda \rightarrow \infty$ .

Notice that the conditions above determine uniquely  $C_j^f$  and  $C_j^g$  for  $j = 0, \dots, \mathcal{N}$  as well as  $f_0$  and  $g_0$ . Namely, if  $\mathcal{N} \geq 3$ , then  $C_j^f$  and  $C_j^g$  are respectively the unique solutions of the systems of linear algebraic equations

$$\begin{aligned} \sum_{j=1}^{\mathcal{N}} C_j^f \int_{\Gamma_j} \frac{\xi^k d\xi}{w(\xi)} &= 0, \quad k = 0, \dots, \mathcal{N} - 2, \\ \sum_{j=1}^{\mathcal{N}} C_j^f \int_{\Gamma_j} \frac{\xi^{\mathcal{N}-1} d\xi}{w(\xi)} &= -2\pi i, \end{aligned} \quad (3.87)$$

and

$$\begin{aligned} \sum_{j=1}^{\mathcal{N}} C_j^g \int_{\Gamma_j} \frac{\xi^k d\xi}{w(\xi)} &= 0, \quad k = 0, \dots, \mathcal{N} - 3, \\ \sum_{j=1}^{\mathcal{N}} C_j^g \int_{\Gamma_j} \frac{\xi^{\mathcal{N}-2} d\xi}{w(\xi)} &= -4\pi i, \\ \sum_{j=1}^{\mathcal{N}} C_j^g \int_{\Gamma_j} \frac{\xi^{\mathcal{N}-1} d\xi}{w(\xi)} &= -2\pi i \sum_{j=0}^{\mathcal{N}} (\lambda_j + \lambda_j^*). \end{aligned} \quad (3.88)$$

If  $\mathcal{N} = 1$ , then  $C_1^f$  and  $C_1^g$  are determined by the last equations in (3.87) and (3.88), if  $\mathcal{N} = 2$ , then  $C_j^f$ ,  $j = 1, 2$  are determined by the system in general form (3.87) whereas  $C_j^g$ ,  $j = 1, 2$  are determined by the system of two last equations in (3.88). Then  $f(\lambda)$  is determined, for all  $\mathcal{N} \geq 1$ , by

$$f(\lambda) = \frac{w(\lambda)}{2\pi i} \sum_{j=1}^{\mathcal{N}} \int_{\Gamma_j} \frac{C_j^f d\xi}{w(\xi)(\xi - \lambda)}, \quad (3.89)$$

whereas  $g(\lambda)$  is determined by

$$g(\lambda) = \frac{w(\lambda)}{2\pi i} \sum_{j=1}^{\mathcal{N}} \int_{\Gamma_j} \frac{C_j^g d\xi}{w(\xi)(\xi - \lambda)} \quad (3.90)$$

for  $\mathcal{N} \geq 2$  and by

$$g(\lambda) = 2w(\lambda) + \frac{w(\lambda)}{2\pi i} \int_{\Gamma_1} \frac{C_1^g d\xi}{w(\xi)(\xi - \lambda)} \quad (3.91)$$

for  $\mathcal{N} = 1$ . In turn,  $f_0$  and  $g_0$  are determined from the asymptotic relations  $f(\lambda) = \lambda + f_0 + O(1/\lambda)$  and  $g(\lambda) = 2\lambda^2 + g_0 + O(1/\lambda)$  as  $\lambda \rightarrow \infty$  of  $f(\lambda)$  and  $g(\lambda)$  already determined above.

It is convenient to introduce the function of two indexes - integer non-negative parameters:  $R(j, k) = \int_{\Gamma_j} \xi^k / w(\xi) d\xi$ . Using this function, one can construct linear equation systems above, eqs. (3.87-3.88) as:

$$R(j, k)C_j^g = \begin{pmatrix} 0 \\ \dots \\ -4\pi i \\ -2\pi i \sum_{j=0}^{\mathcal{N}} (\lambda_j + \lambda_j^*) \end{pmatrix}, \quad (3.92)$$

$$R(j, k)C_j^f = \begin{pmatrix} 0 \\ \dots \\ -2\pi i \end{pmatrix}, \quad k = 0, \dots, \mathcal{N} - 1, \quad j = 1, \dots, \mathcal{N}.$$

In addition, we can fix  $C_0^f = C_0^g = 0$ . Particularly for genus-1 case, the only nonzero  $C_1^{f,g}$  are evaluated as:

$$C_1^f = \frac{-2\pi i}{R(1, 0)}, \quad C_1^g = \frac{-\pi i \sum_{j=0}^1 (\lambda_j + \lambda_j^*)}{R(1, 0)}. \quad (3.93)$$

The additional frequency-like parameters  $f_0$  and  $g_0$  can be given as:

$$f_0 = \frac{1}{2\pi i} \sum_{j=0}^n C_j^f \int_{\Gamma_j} \frac{d\xi}{w(\xi)} W_0(\xi) \quad \text{and} \quad g_0 = \frac{1}{2\pi i} \sum_{j=0}^n C_j^g \int_{\Gamma_j} \frac{d\xi}{w(\xi)} W_0(\xi), \quad (3.94)$$

where  $W_0(\xi)$  is a free term (w.r.t.  $z$ ) of asymptotic series of  $w(z)/(\xi - z)$  at  $z \rightarrow \infty$ , which naturally is a function of  $\xi$ . Expressions above can be also written in terms of the function  $R(j, k)$ .

The solution of RHP  $\mathbf{M}(x, t)$  can be written as  $\mathbf{M} = \mathbf{I} + \hat{C}_\Gamma \mathbf{U}$ , where  $\hat{C}_\Gamma$  is operator, performing Cauchy-type integration along the  $\Gamma = \bigcup_{j=0}^n \Gamma_j$ . Therefore, the corresponding solution of NLS is evaluated as:

$$q(\tau, \zeta) = -\frac{1}{\pi} \int_{\Gamma} \mathbf{U}_{(12)}(x, t, \xi) d\xi e^{2i(f_0\tau + g_0\zeta)}. \quad (3.95)$$

The Riemann-Hilbert solver [107] allows finding directly  $\mathbf{U}(x, t, z)$  and evaluate its integration over the variable  $z$  with the function `DomainIntegrate`.

### 3.3.6 Alternative formulation of RHP in the finite-band case

Solving RHP with jumps on open arcs (or, equivalently, with discontinuous jumps on closed contours) faces the problem of singularities at the arc ends, which affects numerical solutions of the associated integral equations. To remedy this, the original RHP can be transformed to a problem with (continuous) jumps on closed contours. Without loss of generality, I provide here the procedure of deformation for two contours  $\Gamma_j$ , i.e. for  $\mathcal{N} = 1$ .

Introduce two (rectangular in our simulations) contours  $\Xi_{0,1}$  enveloping the arcs  $\Gamma_{0,1}$  with reasonable offset (see fig. 3.7), define the functions:

$$\varkappa_j(\lambda) = \left( \frac{\lambda - \lambda_j}{\lambda - \lambda_j^*} \right)^{1/4} \quad (3.96)$$

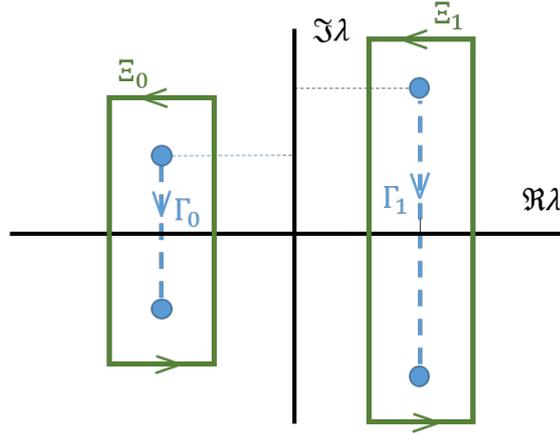


Figure 3.7: Closed contours  $\Xi_{0,1}$ , enveloping the arcs  $\Gamma_{0,1}$  for alternative RHP formulation, which avoids singularity points.

for  $\lambda \in \mathbb{C} \setminus \Gamma_j$  (the branch of the root is fixed by the condition  $\kappa_j(\lambda) \rightarrow 1$  as  $\lambda \rightarrow \infty$ ), and define the matrices  $\mathbf{K}_{0,1}$  by

$$\mathbf{K}_0(\lambda) = \frac{1}{2} \begin{pmatrix} \varkappa_0(\lambda) + \frac{1}{\varkappa_0(\lambda)} & -\varkappa_0(\lambda) + \frac{1}{\varkappa_0(\lambda)} \\ -\varkappa_0(\lambda) + \frac{1}{\varkappa_0(\lambda)} & \varkappa_0(\lambda) + \frac{1}{\varkappa_0(\lambda)} \end{pmatrix} \quad (3.97)$$

and

$$\mathbf{K}_1(\tau, \zeta, \lambda) = \frac{1}{2} \begin{pmatrix} \varkappa_1(\lambda) + \frac{1}{\varkappa_1(\lambda)} & \frac{\varkappa_1(\lambda) - \frac{1}{\varkappa_1(\lambda)}}{i\theta(\tau, \zeta)} \\ i\theta(\tau, \zeta) \left( \varkappa_1(\lambda) - \frac{1}{\varkappa_1(\lambda)} \right) & \varkappa_1(\lambda) + \frac{1}{\varkappa_1(\lambda)} \end{pmatrix}. \quad (3.98)$$

Introduce  $\hat{\mathbf{M}}$  by

$$\hat{\mathbf{M}} = \begin{cases} \mathbf{M}\mathbf{K}_j^{-1}, & \text{inside } \Xi_j, \\ \mathbf{M}, & \text{outside } \Xi_j, \end{cases} \quad (3.99)$$

where  $\mathbf{M}$  is the solution of the original RHP (3.85). Then  $\hat{\mathbf{M}}$  satisfies RHP with jump continuous on each closed part  $\Xi_j$  of the total contour:

$$\begin{aligned} \hat{\mathbf{M}}^-(\tau, \zeta\lambda) &= \hat{\mathbf{M}}^+(\tau, \zeta\lambda)\mathbf{K}_j, \quad \lambda \in \Xi_j, \quad j = 0, 1, \\ \hat{\mathbf{M}}(\tau, \zeta, \lambda \rightarrow \infty) &\rightarrow \mathbf{I} \end{aligned} \quad (3.100)$$

(notice that  $\hat{\mathbf{M}}$ , in contrast to  $\mathbf{M}$ , has no jumps across  $\Gamma_{0,1}$ ). Such a problem can be efficiently solved numerically using the RHP solver [107]. On the other hand, the solution  $\hat{\mathbf{M}}$  of the RHP (3.100) still gives rise to the genus-1 solution of the NLS by  $q(\tau, \zeta) = 2i(\hat{\mathbf{M}}^{(i)})_{(12)}(\tau, \zeta)$ .

The presented RHP deformation makes the numerical computation more reliable for further communication applications.

### 3.3.7 Algebraic-geometric approach

Appart form the RHP approach, finite-genus solutions of NLS, eq. (2.15) are conventionally expressed in terms of meromorphic functions, differentials and integrals on hyperelliptic Riemann surfaces of finite genus, see [14]. The construction of such solutions (usually called in this context *algebraic-geometric*) involves the following steps:

- (i) Given  $\{\lambda_j\}$ ,  $j = 0, 1, \dots, \mathcal{N}$ , introduce the Riemann surface associated with the algebraic curve  $F(w, \lambda) = 0$  having the form

$$w^2 = \prod_{j=0}^{\mathcal{N}} (\lambda - \lambda_j)(\lambda - \lambda_j^*).$$

This is a common starting point for both RHP and algebro-geometric approach, however, we do not treat  $w(z)$  as a seed for the Riemann surface.

- (ii) Introduce the canonical basis of oriented circles  $(a_j, b_j)$ ,  $j = 1, \dots, \mathcal{N}$ , on the above Riemann surface, and define the associated objects: the period matrix  $\mathbf{B}$ , the basis of holomorphic differentials  $\boldsymbol{\omega}$ , and the Abel integrals  $\Omega_j$ ,  $j = 1, 2, 3$ , fixed by their particular behaviour at  $\infty^\pm$  (the infinity points at two sheets of the Riemann surface).
- (iii) Determine the scalar parameters  $E$ ,  $N$ , and  $\alpha$  from the asymptotic expansions of  $\Omega_j$  at  $\infty^\pm$ , and the vector parameters  $\mathbf{V}$ ,  $\mathbf{W}$ , and  $\mathbf{r}$  by  $V_j = \int_{b_j} d\Omega_1$ ,  $W_j = \int_{b_j} d\Omega_2$ ,  $j = 1, \dots, \mathcal{N}$ ,  $\mathbf{r} = \int_{\infty^-}^{\infty^+} \boldsymbol{\omega}$ .
- (iv) Given  $\mathbf{B}$ , determine the Riemann theta function

$$\Theta(u_1, \dots, u_{\mathcal{N}}) = \sum_{l \in \mathbb{Z}^{\mathcal{N}}} \exp \left[ \frac{1}{2} (\mathbf{B} l, l) + (l, u) \right], \quad (3.101)$$

where  $(l, u) = l_1 u_1 + \dots + l_{\mathcal{N}} u_{\mathcal{N}}$  (similar to dot product).

- (iv) Given an additional  $\mathcal{N}$ -dimensional vector  $\mathbf{D}$ , determine the NLS solution  $q(\tau, \zeta)$  by

$$q(t, z) = \alpha \frac{\Theta(i\mathbf{V}\tau + i\mathbf{W}\zeta - \mathbf{D} + \mathbf{r})}{\Theta(i\mathbf{V}\tau + i\mathbf{W}\zeta - \mathbf{D})} \frac{\Theta(\mathbf{D})}{\Theta(\mathbf{D} - \mathbf{r})} e^{-if_0\tau + ig_0\zeta}. \quad (3.102)$$

In the procedure above, the auxiliary matrix and scalar objects are determined by the following conditions.

Functions  $f(\lambda)$  and  $g(\lambda)$  defined by eq. (3.84) can be expressed via hyperelliptic integrals on  $\mathbb{C} \setminus \Gamma$ . Period matrix  $\mathbf{B}$  is defined through the normalised holomorphic differentials as  $B_{(jk)} = \int_{b_k} d\omega_j$ . Here  $b_k$ -cycle starts from the left arc  $(\lambda_0, \lambda_0^*)$ , goes on the upper sheet of the Riemann surface to  $(\lambda_k, \lambda_k^*)$ , and returns on the lower sheet to the starting point. This is a symmetric matrix with positive definite imaginary part. The Abelian integrals are given as:

$$\omega_j = \int_{\lambda_0}^z \psi_j(s) ds, \quad j = 1, 2, \dots, \mathcal{N}, \quad (3.103)$$

so the  $d\omega_j$  are basis holomorphic differentials on the Riemann surface:

$$\psi_j(z) = \frac{\sum_{m=1}^{\mathcal{N}} c_{jm} z^{\mathcal{N}-m}}{w(z)}. \quad (3.104)$$

The highest importance here is how the parameters  $\tau, \zeta$  enter the expression eq. (3.101).  $\mathbf{V}$  and  $\mathbf{W}$  are linked with frequency and wavenumbers from eq. (3.92):

$$V_{(k)} \sim -\frac{1}{2\pi} C_k^f, \quad W_{(k)} \sim -\frac{1}{2\pi} C_k^g. \quad (3.105)$$

Finally, the norming coefficient  $\alpha$  is defined through the main spectrum points as  $\alpha = 1/4 \sum_{j=0}^{\mathcal{N}} (\lambda_j + \lambda_j^*)$ .

Although the construction of  $q(\tau, \zeta)$  above can be considered as explicit, its numerical implementation faces several drawbacks, one of them being related to the numerical

calculation of Riemann theta function. The first step in the calculation of  $\Theta$  as a multi-dimensional Fourier series is to truncate the grid from which the vectors  $\mathbf{l}$  are drawn, to a bounded subspace of  $\mathbb{C}^{\mathcal{N}}$ , thus introducing the sampling of optical domain. This itself introduces some error in the calculations, which can be controlled to some extent by considering the contribution of different terms along with the series [29]. However, if this bounded grid is limited to integers between  $-M_1$  and  $M_1$ , the number of terms in (3.101) is  $(2M_1 + 1)^{\mathcal{N}}$ . It means that even for the simple case of  $\mathcal{N} = 2$ , when we have just three  $\lambda_i$  to encode our information (genus-2 solution), we gain the quadratic growth of the number of elementary operations (with the number of samples  $M_1$ ) required for the computation of respective theta-function, and this is already higher than the complexity of fast inverse NFT methods [142].

On the other hand, the argument of Riemann theta function in eq. (3.101), contains a literally linear dependency on time. It motivates that we can treat the parameters  $C^{f,g}$  from eqs. (3.92), (3.105) as frequencies and wavenumbers, respectively. Taking this into account, we can ensure the periodicity using a commensurable set of frequencies,  $C_j^f$ , to convert the multi-dimensional series into an ordinary Fourier series. In this way, calculating the Riemann theta function is turned into a Fourier series with time-dependent coefficients coming from another Fourier series. Although the latter property helps decrease the computational complexity, it remains considerably heavy, especially when compared to its alternative that is rendered by the RHP approach.

### 3.3.8 Inverse problem and signal generation discussion

Above we presented several approaches to the inverse NFT. From the practical perspective in the application for communication, the inverse part determines the entire structure of the transmission system, when the forward one acts more complimentary. First, because of the distribution over the transceiver parts: inverse NFT is performed on the Tx, when forward NFT - on the Rx. Therefore, the chosen way of INFT defines the modulation opportunities, and the anticipated signal parameters.

In the design of the transmission system, apart from satisfying mathematical expressions to make the IST concept work, we have to step apart from the pure equational interpretation and get the physical understanding. NLS as a model of the signal propagation in the fibre is derived under certain assumptions, so the generated signals are practically implementable only when they meet these assumptions. However, because the limitations bound the outcome signals, but not their NF representation, it becomes a challenging part of the communication system design process. For instance, that puts us much beyond well-studied cases of non-zero background IST, like step-shape background, because its asymptotes at infinities do not coincide.

I want to emphasize here that the limitations arise exactly at the stage of appropriate inverse NFT choice, together with the shaping of the NF domain picture. It is a decision point, which determines the entire format and performance of the transmission. What we care about the most are common signal properties: power, bandwidth, duration of the informative part.

In the vanishing case, there are more theoretical results, which facilitate the design of the signal with predictable properties. The explicit correspondence between the discrete spectrum subparts and the related soliton parameters (subsection 3.1.3) provides some clues about the outcome signal structure, therefore, the pure eigenvalue modulation is the most attractive from this perspective. Additionally, the Darboux transform, especially its fast numerical realisation [137], makes this entire concept affordable, however, the poor opportunities in the discrete spectrum for data modulation makes the developing of this approach meaningless on the way of reaching the ultimate goal of nonlinearity mitigation - increasing the channel capacity.

However, all other kinds of modulation are not so straightforward from the point of controlling and managing signal parameters. Nonlinear Parseval's identity figures out the outcome signal energy from the NF spectrum energy, but it is not sufficient for full control. We lack the analogy of Nyquist sampling theorem.

Another challenging point is the numerical error arisen from the algorithms. It is the most severe degradation factor, especially it the Tx (inverse NFT) side. First, it is inevitable because the mathematical relations, which are used for algorithms foundations, are provided for continuous waveforms of the signal and the NF spectrum, when all computations are done for their sampled versions. In short, we do not have a proper analogy of discrete FT, which it perfectly operates back-to-back up to a single sample. The GLM numerical method, one of the most fundamental, operate in a sample-by-sample way, so, it accumulates the error, gained at each step. Some benefits we can get from the RHP approach, because it processes all samples independently, however, the resulting signals, as found in this study, have a noticeable impact of truncations, so cannot be used for the signal generation.

In the finite-band case, there two approaches presented. Despite the algebro-geometric one provides an explicit expression for a finite-gap NLS solution, it is impractical for the computation, especially when we design a system with the high prospective data rate. The RHP method provides linear scaling of the complexity with a number of bands, and more flexibility of the mathematical problem formulation on the complex plane. The problem of the signal parameters control still has a place here, similarly to the vanishing case, but at least the signal duration and bandwidth are explicitly linked to the main spectrum structure up to the first-order accuracy (scaling with the signal power).

The general concept of application IST to the modulation formats for mitigating non-linearity in optical communication requires the accurate following the mathematical assumptions, their linkage with real physics properties, controlling and possibly reducing the numerical error, and having sufficient degrees of freedom for data carrying for providing effective data rate. The choice of the appropriate inverse NFT method, combined with the determination of flexibilities and constrains in the NF image, is a crucial decision-making part of the entire transmission system design.

## Chapter 4

# Numerical methods for scattering data evaluation

For the effective application of NFT technique for nonlinearity mitigation in the optical fibres, it is crucial to use sufficiently accurate and fast numerical algorithms. This chapter is dedicated to the review of the numerical methods for direct NFT and analyse of their numerical accuracy, performance and runtime for computation of the continuous and discrete spectral data, see section 3.1. The results in this chapter were published as a part of the review paper [J2] and partly reported in the conference publications [C8,C9]. In this chapter, I mainly focus on real-valued  $\lambda$  and use vanishing boundary conditions case and the right set of scattering data  $\{r(\lambda), c_j\}$  to test the performance of the algorithms. However, the results can be generalised to the periodic case, because in both cases the direct transform is done via ZSP solution.

### 4.1 Test profiles and metrics

The presented algorithms are tested against three independent test profiles. To ensure maximum possible variation, I involve wide-known model signals, where the analytical expressions for the scattering data can be written explicitly.

- (i) The over-soliton potential [126] (or Satsuma-Yajima pulses) is given by

$$q_{\text{over}}(\tau) = A \operatorname{sech} \tau. \quad (4.1)$$

It is characterized by a single real amplitude parameter  $A > 0$ . The associated spectral functions are as follows (all quantities with subscript over represent over-solitons):

$$a_{\text{over}}(\lambda) = \frac{\Gamma^2\left(\frac{1}{2} - i\lambda\right)}{\Gamma\left(-A - i\lambda + \frac{1}{2}\right) \Gamma\left(A - i\lambda + \frac{1}{2}\right)}, \quad (4.2)$$

$$b_{\text{over}}(\lambda) = -\sin(\pi A) \operatorname{sech}(\pi\lambda), \quad (4.3)$$

and

$$r_{\text{over}}(\lambda) = -\frac{\sin(\pi A) \operatorname{sech}(\pi\lambda) \Gamma\left(-A - i\lambda + \frac{1}{2}\right) \Gamma\left(A - i\lambda + \frac{1}{2}\right)}{\Gamma^2\left(\frac{1}{2} - i\lambda\right)}, \quad (4.4)$$

where  $\Gamma(\cdot)$  is the Euler Gamma function.

Depending on the value of  $A$ , the discrete spectrum attributed to the over-soliton (4.1) consists of simple eigenvalues

$$\lambda_k = (A - 1/2 - k)i, \quad k = 0 \dots \left\lfloor A - \frac{1}{2} \right\rfloor. \quad (4.5)$$

If  $A$  is exactly half-integer, then  $r(\lambda) = 0$  and the total energy is completely concentrated in the solitonic modes. The norming constant corresponding to the highest eigenvalue  $\lambda_0 = (A - 1/2)i$  is

$$c_{\text{cover}} = i\Gamma(2A)/\Gamma^2(A). \quad (4.6)$$

More details on the NFT properties of the profiles (4.1) can be found in [126, 133]. Similar profile was used for verification of RHP approach for vanishing case, see subsection 3.3.4.

(ii) The ZSP for the rectangular potential

$$q_{\text{rec}}(\tau) = \begin{cases} A, & -L \leq \tau \leq L \\ 0, & \text{otherwise} \end{cases} \quad (4.7)$$

can also be solved analytically [18, 19]. The associated scattering coefficients are given by

$$a_{\text{rec}}(\lambda) = e^{2i\lambda L} \left( \cos \left[ 2\sqrt{\lambda^2 + A^2} L \right] - \frac{i\lambda}{\sqrt{\lambda^2 + A^2}} \sin \left[ 2\sqrt{\lambda^2 + A^2} L \right] \right), \quad (4.8)$$

$$b_{\text{rec}}(\lambda) = \frac{A}{\sqrt{\lambda^2 + A^2}} \sin \left[ 2\sqrt{\lambda^2 + A^2} L \right], \quad (4.9)$$

and

$$r_{\text{rec}}(\lambda) = \frac{A \exp(-2i\lambda L)}{i\xi - \sqrt{\lambda^2 + A^2} \cot \left[ 2\sqrt{\lambda^2 + A^2} L \right]}. \quad (4.10)$$

The discrete eigenvalues  $\{\lambda_{\text{rec}}\}$  for the rectangle profile are given by the roots of the following transcendental equation:

$$\tan \left[ 2\sqrt{A^2 + \lambda_{\text{rec}}^2} L \right] = \frac{\sqrt{A^2 + \lambda_{\text{rec}}^2}}{i\lambda_{\text{rec}}} \quad (4.11)$$

for  $\lambda_{\text{rec}}$  in the upper half-plane. The norming constant for  $\lambda_{\text{rec}}$  is given by the expression

$$c_{\text{rec}}(\lambda) = -\frac{i(A^2 + \lambda^2) e^{-2iL\lambda}}{A \left( 2L\sqrt{A^2 + \lambda^2} \cot \left[ 2\sqrt{A^2 + \lambda^2} L \right] - 1 \right)}. \quad (4.12)$$

(iii) For our tests, we have also used the solitonic potential ( $r(\lambda) = 0$ ,  $\lambda \in \mathbb{R}$ ) with unit amplitude and phase [19]:

$$q_{\text{sol}}(\tau) = \exp(-i\tau) \text{sech}(\tau). \quad (4.13)$$

It has a single eigenvalue  $\lambda_{\text{sol}} = 0.5 + 0.5i$  with the associated norming constant  $c_{\text{sol}} = i$ . This potential allows us to check the behaviour of our methods in the case of eigenvalues having a non-zero real part.

As metrics of the numerical accuracy, I involve the (mean) relative error. For discrete  $y$ , I use relative error as an accuracy descriptor:

$$\epsilon = \frac{|y^{(\text{computed})} - y^{(\text{analytical})}|}{|y^{(\text{analytical})}|}. \quad (4.14)$$

For a continuous spectral function  $f(\lambda)$  ( $f_i$  can be replaced for  $a$ ,  $b$  or  $r$ ), I compare the analytical and computed values using the mean squared relative error (MSRE):

$$\epsilon_f = \frac{1}{N} \sum_{k=1}^N \frac{|f_k^{(\text{computed})} - f_k^{(\text{analytical})}|^2}{|f_k^{(\text{analytical})}|^2}, \quad (4.15)$$

where  $f_k = f(\lambda_k)$ . For  $f^{(\text{analytical})}(\xi) = 0$ , I use the ordinary squared difference  $[|f_k^{(\text{computed})} - f_k^{(\text{analytical})}|^2]$ .

Here I have to emphasise the principal difference between above-listed profiles and practical signals used in optical communications. The analytical profiles are rather smooth and simply structured, when in communications, we modulate the signal from the random data stream, using various modulation approaches, see subsection 2.6.1. In particular, for full usage of the transmission opportunities, both time window and bandwidth are occupied to the maximum. It means that the resulting signals are complex and highly oscillating, which makes the verification of the algorithms on the base of smooth profiles less valuable. However, it is a common practice in numerical NFT community [23, 38, 101, 102, 103, 141], so in this research, I followed this tradition too. In addition, I want to note that this chapter is based on my publication, done mainly in 2017. From this time several higher-order methods have been developed, utilising exponential factorisation and non-uniform sampling [23, 101, 102].

## 4.2 Transfer-matrix methods

The transfer-matrix approach is the common name of the family of algorithms, which use a certain truncation and sampling of ZSP, considering the potential being constant through a single step.

For the truncation of model potentials and discretization,  $q(\tau)$  is analytically represented to define  $N_q$  samples within the interval  $\tau \in [-L, L]$ . The interval is divided into  $n$  equal subintervals of length  $\Delta\tau = 2L/N_q$ , where the  $m$ -th subinterval is  $\tau \in [\tau_m - \Delta\tau/2, \tau_m + \Delta\tau/2]$ . Outside of the interval  $[-L, L]$ , the potential is assumed to be exactly zero. For all methods considered below, our signal (ZSP potential) is approximated by a constant value along a single step:  $q_m = q(\tau_m)$ .

The vector of wave functions  $\Psi(\tau, \lambda) = (\psi_1(\tau, \lambda), \psi_2(\tau, \lambda))^T$  from eq. (3.5), is fixed by imposing the initial conditions on the left edge of the truncation interval ( $\tau = -L$ ): according left-to-right wave definition  $\psi^{lr}(\tau, \lambda)$  from eq. (3.12),  $\Psi(-L) = (e^{i\lambda L}, 0)^T$ . The vector of envelopes  $\mathbf{X}(\tau) = (\chi_1(\tau), \chi_2(\tau))^T$  from eq. (3.27) is fixed correspondingly as  $\mathbf{X}(-L) = (1, 0)^T$ . The spectral functions  $a(\lambda)$  and  $b(\lambda)$  are defined on the right edge of the processing interval ( $\tau = L$ ): according to (3.17) and (3.28),  $a(\lambda) = \psi_1(L)e^{i\lambda L}$  and  $b(\lambda) = \psi_2(L)e^{-i\lambda L}$ , or, alternatively,  $a(\lambda) = \chi_1(L)$  and  $b(\lambda) = \chi_2(L)$ .

To propagate the incident wave towards the end of the processing interval, many approaches use the transfer matrix  $T_m$  for propagating ZSP solution over a single  $m$ -th discretization step, inside which the potential is considered as a constant,  $q_m = \text{const}$ :

$$\Psi_{m+1} = \mathbf{T}_m \Psi_m. \quad (4.16)$$

Performing the consequent iterations of eq. (4.16) from  $m = 1$  to  $m = N_q$ , we eventually find the desired values of the Jost functions at the end of the processing interval and compute the NF parameters  $a(\lambda)$  and  $b(\lambda)$ .

- (i) For Bofetta-Osborne (BO) method [18, 19], we define  $\mathbf{T}_m$  evaluating the matrix exponential of the matrix in ZSP with constant potential  $q_m$ :

$$\mathbf{T}_m^{(\text{BO})} = \exp \left[ \Delta\tau \begin{pmatrix} -i\lambda & q_m \\ -q_m^* & i\lambda \end{pmatrix} \right], \quad (4.17)$$

which can be evaluated explicitly:

$$\mathbf{T}_m^{(\text{BO})} = \begin{pmatrix} \cosh \kappa \Delta \tau - i\lambda/\kappa \sinh \kappa \Delta \tau & q_m/\kappa \sinh \kappa \Delta \tau \\ -q_m^*/\kappa \sinh \kappa \Delta \tau & \cosh \kappa \Delta \tau + i\lambda/\kappa \sinh \kappa \Delta \tau \end{pmatrix}, \quad (4.18)$$

with  $\kappa = \sqrt{-|q_m|^2 - \lambda^2}$ .

- (ii) For Ablowitz-Ladik (AL) method, we use the normalized discretization of ZSP [140, 141, 151] and apply Euler method, then substituting  $1 \pm i\lambda\Delta\tau$  with  $\exp(\pm i\lambda\Delta\tau)$ , we get the transfer matrix:

$$\mathbf{T}_m^{(\text{AL})} = \frac{1}{\sqrt{1 + \Delta\tau^2|q_m|^2}} \begin{pmatrix} e^{-i\lambda\Delta\tau} & \Delta\tau q_m \\ -\Delta\tau q_m^* & e^{i\lambda\Delta\tau} \end{pmatrix}. \quad (4.19)$$

The AL method with the norming factor  $1/\sqrt{1 + \Delta\tau^2|q_m|^2}$  provides higher stability and accuracy compared to that without the normalisation [140].

In this research, I also derived and studies two novel modified versions of BO and AL algorithms, for which the similar approaches are applied to ZSP for the envelope functions (3.27). Evolution over each step  $\Delta\tau$  is again performed using the transfer matrices:

$$\mathbf{X}_{m+1} = \mathbf{T}_m \mathbf{X}_m. \quad (4.20)$$

- (iii) For the modified BO method, applying the matrix exponential to the matrix of ZSP for the envelope functions defined in (3.27), one can get the transfer matrix in the form:

$$\mathbf{T}_m^{(\text{BOmod})} = \begin{pmatrix} \cos |q_m \Delta \tau| & \sin |q_m \Delta \tau| e^{i(\theta_{q_m} + 2\lambda\tau)} \\ -\sin |q_m \Delta \tau| e^{-i(\theta_{q_m} + 2\lambda\tau)} & \cos |q_m \Delta \tau| \end{pmatrix}, \quad (4.21)$$

where  $\theta_{q_m}$  is  $\arg[q_m]$ .

- (iv) In the case of the modified AL method, we have

$$\mathbf{T}_m^{(\text{ALmod})} = \frac{1}{\sqrt{1 + \Delta\tau^2|q_m|^2}} \begin{pmatrix} 1 & q_m \Delta \tau e^{2i\lambda\tau} \\ -q_m^* \Delta \tau e^{-2i\lambda\tau} & 1 \end{pmatrix}. \quad (4.22)$$

- (v) Finally, for the (non-modified) Crank-Nikolson (CN) method [140, 151], the transfer matrix entering eq. (4.16) is given by

$$\mathbf{T}_m^{(\text{CN})} = (\mathbf{I} - \frac{\Delta\tau}{2} \mathbf{P}_{m+1})^{-1} (\mathbf{I} + \frac{\Delta\tau}{2} \mathbf{P}_m), \quad (4.23)$$

where

$$\mathbf{P}_m = \begin{pmatrix} -i\lambda & q_m \\ -q_m^* & i\lambda \end{pmatrix}. \quad (4.24)$$

Notice that in the NFT related works (see e.g. [19, 151]) the algorithms for the solutions of ordinary differential equation (the Runge-Kutta scheme) were also studied in the application of ZSP analysis. The Runge-Kutta fourth-order algorithm (RK) for the solution of ZSP is better applicable to the envelope system, eq. (3.27), as in this case the rapid oscillations of the Jost functions (for the region of  $\lambda$  with a large real part) are included into the effective potential functions. We do not describe the RK scheme here as it is quite standard (see e.g. [19]) but below we present a comparative analysis of the RK algorithm with the algorithms mentioned above.



Figure 4.1: Transparency scale for amplitude variation in grey tones; progressively darker shades indicate higher amplitudes

Now we perform a comparative accuracy analysis of the described methods using the three scattering functions  $a(\lambda)$ ,  $b(\lambda)$ , and (recall that we use right scattering function)  $r(\lambda)$ , eqs. (3.17-3.18). For the real values of  $\lambda$  the MSRE (4.15) is used for the accuracy assessment.

We also address the behaviour of NFT methods in analysing the dependence of the method's accuracy on the variation of amplitude: in all the following figures, the amplitude changes are depicted using the transparency scale of the corresponding curves, see fig. 4.1, i.e. the curves for different amplitudes are plotted by the coloured areas changing from an almost transparent part (the lowest amplitudes) to an almost solid line (the highest amplitude). Captions to the plots provide information about the range and step of amplitude variation.

Our analysis confirms that the BO method gives the best accuracy among all methods studied, see fig. 4.2. The AL and the modified BO algorithms display similar behaviour with the change of amplitude  $A$  and of the number of points  $N_q$ , whereas the CN algorithm came up with the worst accuracy and convergence rate. Both AL and BO methods have the same convergence (inclination of the curves on logarithmic scale plots), implying that they all have the same order of accuracy. This conclusion complies with the results summarized in [135] and earlier studies. As expected, the fourth-order RK method converges faster. This method can be better than BO for big enough values of  $N_q$ . At the same time, for a smaller number of discretization points, the RK method's MSRE is excessively large. The BO method shows the weakest dependence on the amplitude variation, whereas the RK is the most sensitive to it (the error increment can reach several orders of magnitude in the range of amplitudes that we used for our plots). We also note that the BO method in application to the rectangular potential gives the solution, which coincides with an explicit analytical expression, so that the main source of errors here is the computational error in evaluation of cosh and sinh from (4.18). This offsets the surprising increase in error with an increase in the number of points for this particular potential and method.

The energy embedded into the continuous NF spectrum, eq. (3.22) can also be a convenient quantity for methods' accuracy assessment: in particular, it was used as a qualifying metric in [19]. We analyse the relative error in the calculation of energy (4.14) versus the number of points and amplitude, see fig. 4.3.

The energy analysis shows the qualitatively similar result as the MSRE analysis does, which confirms its correctness for the NFT methods' accuracy assessment. The RK algorithm converges more rapidly than all other methods, especially for the pure solitonic potential, where the double-precision numbers are not enough to find out the difference between the analytical and computed energy values. At the same time, this test reveals yet another disadvantage of the RK method: it is extremely slow in comparison with all other methods (see table 4.1). The second-worst in terms of time consumption is CN method, whereas both AL and the modified BO are similar in terms of computational time consumption and are the fastest among the all discussed methods. These methods are approximately two times faster than the ordinary BO algorithm, but the latter has impressively high accuracy. The results for the amplitude dependence when using the

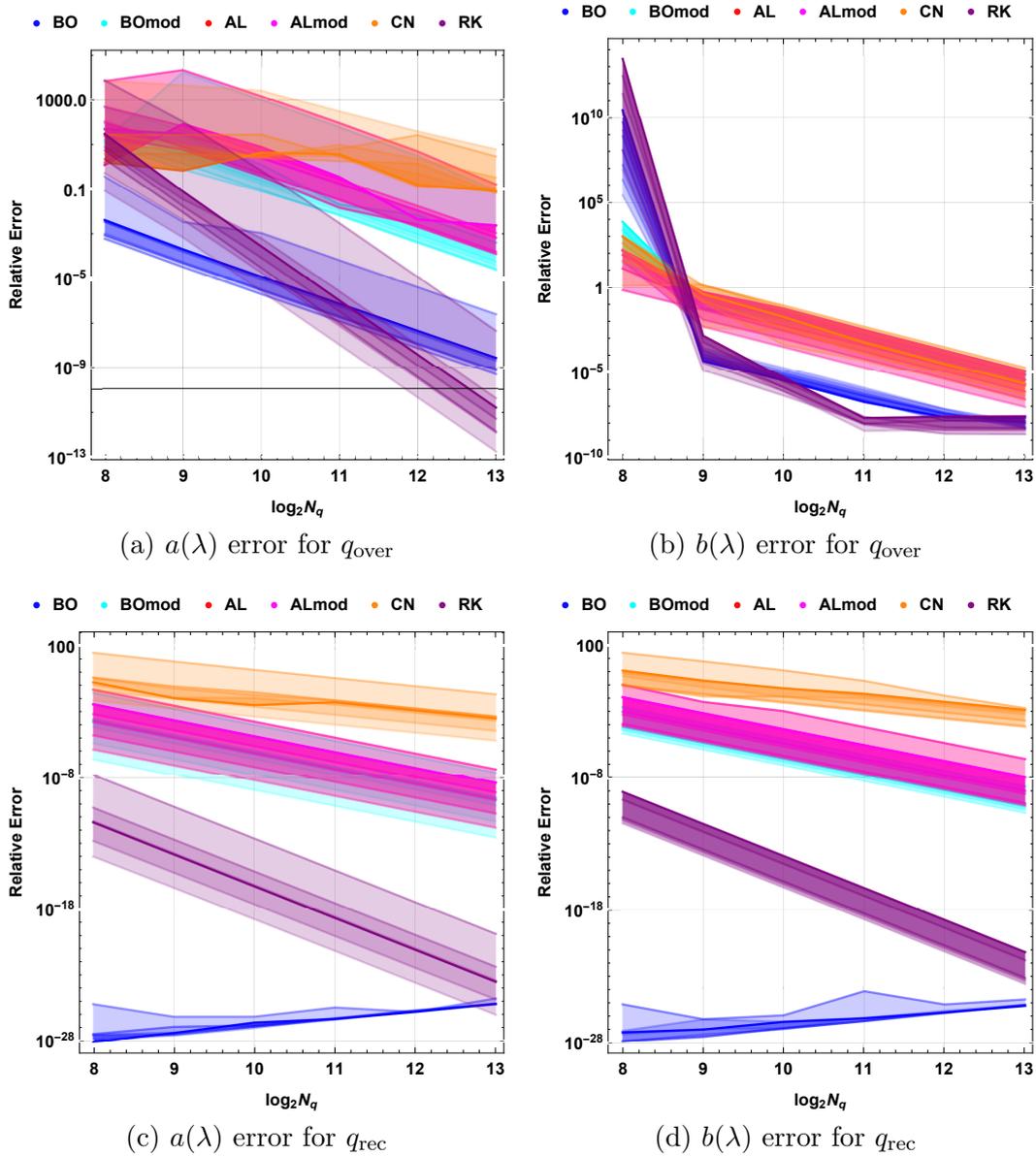


Figure 4.2: MSRE for the computation of the NF scattering functions as a function of the number of discretization points  $N_q$ , evaluated via different NFT methods and for different test profiles amplitudes  $A$ . For  $q_{over}$ :  $A \in [2.25, 5.25]$ ,  $\Delta A = 0.5$ ,  $L = 30$ ; for  $q_{rec}$ :  $A \in [2, 5]$ ,  $\Delta A = 0.5$ ,  $L = 1$ .

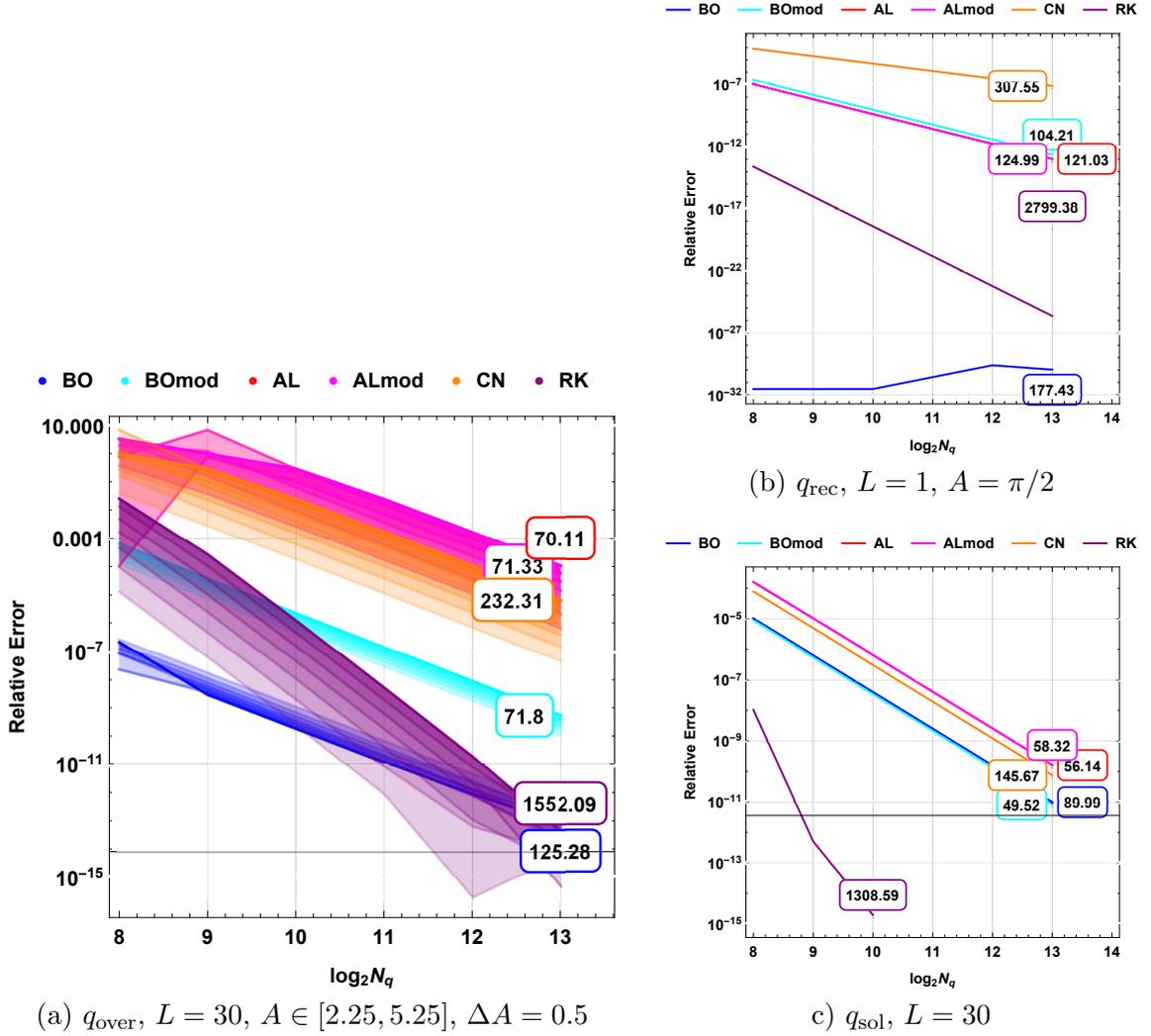


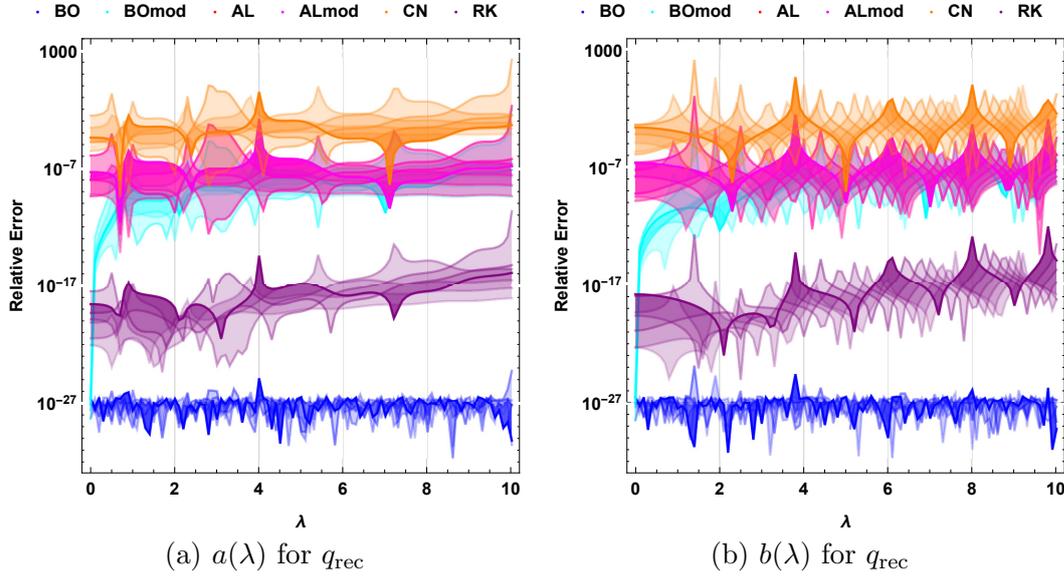
Figure 4.3: Relative error of the continuous spectrum energy  $\varepsilon_q$  (3.22) as a function of number of discretization points  $N_q$  and signal amplitude  $A$  for different NFT methods and different test potentials.

energy as a metric, fig. 4.3, are similar to our findings when the MSRE for NF continuous data was used, fig. 4.2.

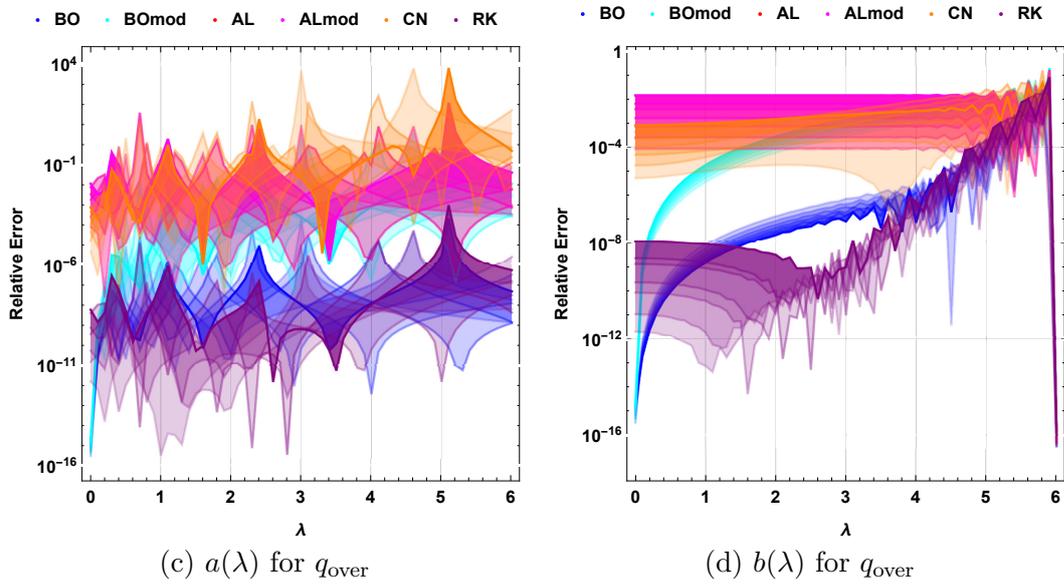
In order to analyse the NFT methods stability in dependence on the nonlinear frequency bandwidth, we investigate the accuracy of our methods along the nonlinear frequency  $\lambda$  axis point-by-point, see fig. 4.4.

As concluded before, BO and RK methods typically show higher integral accuracy. In estimating the accuracy dependence of the frequency, we observe that the numerical error deviates dramatically for the above-mentioned methods, while for AL, ALmod, BMod and CN methods the numerical error is a lot more stable. This effect is more pronounced for the over-soliton potential. This tendency does not change significantly when we are tuning the amplitude of the potential. The CN method demonstrates the worth accuracy among the all methods studied.

Now let us compare our findings with the previous results. The accuracy assessment of computing the continuous spectral data was done in [19]. The BO and RK methods were compared there with regard to the continuous spectrum energy computation: the convergence of the methods was studied and their runtimes were analyzed. It was shown that for smooth solitonic potentials, the RK method was better than the BO method, but



$$L = 1, A \in [2, 5], \Delta A = 0.5$$



$$L = 20, A \in [2.25, 5.25], \Delta A = 0.5$$

Figure 4.4: Relative error of the NF scattering functions evaluation as a function of spectral parameter  $\lambda$ ,  $\Delta\lambda = 0.1$ , and amplitude  $A$  ( $N_q = 2^{10}$ ).

Table 4.1: Runtimes (in seconds) of continuous energy evaluation for different NFT algorithms ( $N_q = 2^{13}$ ).

NFT method					
BO	BOmod	AL	ALmod	CN	RK
$q_{\text{over}}, L = 30, A = 5.25$					
125.28	71.8	70.11	71.33	232.31	1552.99
$q_{\text{rec}}, L = 1, A = \pi/2$					
177.43	104.21	121.03	124.99	307.55	2799.38
$q_{\text{sol}}, L = 30$					
89.99	49.52	56.14	58.32	145.67	1308.59

the authors [19] attributed this finding to the properties of the CPU architecture used for their computations. For the rectangular potential, which has sharp edges, the RK method was shown to perform slower than the BO one. We have generally observed that in terms of runtime and resulting accuracy, the BO method typically outperforms the other options; we have also noticed the excessively large runtime required by the RK methods in comparison to any transfer-matrix based approach of the type shown in eq. (4.16). The authors of [140] and [151] presented the detailed description of various NFT algorithms, but they were mainly focused on the discrete eigenvalues computation accuracy. The authors of [140] also proposed the fast implementation of the AL algorithms based on the FFT-type ideology for the matrix product computation and multipoint fast polynomial evaluation, and this study was continued in [141] for the periodic potential. The comparison of the *conventional* non-fast NFT methods accuracy for the case of periodic potentials and discrete eigenvalues were also presented in [63]. Typically, the qualitative behaviour of the accuracy as a function of  $N_q$  followed the scenario described above in this section.

#### 4.2.1 Derivative computation

The methods of the discrete spectrum components evaluation (iterative from section 4.3 or contour integrals from chapter 5) may require us to find the value of  $a'(\lambda)$  together with the value of  $a(\lambda)$  at the same point  $\lambda$ . Since function  $a(\lambda)$  is homomorphic in the upper half-plane of  $\lambda$  (see, for example, [2]), the divided difference can be used to approximate the derivative. However, it is also possible to find the value of the derivative more accurately in the same programming loop together with the function computation itself. A similar approach is described in [18, 19, 51, 151]. The idea here is to evolve the derivative of  $\Psi'(\tau, \lambda)$  along  $\tau$  processing interval together with  $\Psi(\tau, \lambda)$  itself. In the case of ordinary ZSP (3.5), the evolution starts with the initial condition for the derivative defined at  $\tau \rightarrow -\infty$ :  $\Psi'(\tau \rightarrow -\infty, \lambda) \rightarrow (-i\tau e^{-i\lambda\tau}, 0)^T$ . After the truncation of the  $\tau$ -interval, the initial condition takes the form  $\Psi'(-L, \lambda) = (iLe^{i\lambda L}, 0)^T$ . For the ZSP written for the envelope function, eq. (3.27), we have  $\mathbf{X}'(\tau \rightarrow -\infty, \lambda) = (0, 0)^T$ , or, after the truncation,  $\mathbf{X}'(-L, \lambda) = (0, 0)^T$ . Evolution of the wave function derivative over a single step is performed by applying the relation:

$$\Psi'_{m+1} = \mathbf{T}'_m \Psi_m + \mathbf{T}_m \Psi'_m, \quad \text{or} \quad \mathbf{X}'_{m+1} = \mathbf{T}'_m \mathbf{X}_m + \mathbf{T}_m \mathbf{X}'_m. \quad (4.25)$$

Thus, for all transfer matrix methods, the matrix  $\mathbf{T}'_m$  can be easily determined. At the same time, the RK method does not have that advantage: here we have to numerically

solve the ZSP for the derivatives separately, equipped with different initial conditions.

Our algorithm requires the use of the derivatives of the transfer matrices of each NFT algorithm described above. In the case of the BO matrix, the calculation of the derivative yields:

$$\mathbf{T}'_m^{(\text{BO})} = \begin{pmatrix} A_+ & B \\ C & A_- \end{pmatrix}, \quad (4.26)$$

where

$$\begin{aligned} A_{\pm} &= \pm \frac{i\lambda^2 \Delta\tau}{\kappa^2} \cosh \kappa \Delta\tau \mp \sinh \kappa \Delta\tau \left( \pm \frac{\lambda \Delta\tau}{\kappa} + \frac{i}{\kappa} + \frac{i\lambda^2}{\kappa^3} \right), \\ B &= \frac{q_m \lambda}{\kappa^3} \sinh \kappa \Delta\tau - \frac{q_m \lambda \Delta\tau}{\kappa^2} \cosh \kappa \Delta\tau, \\ C &= -\frac{q_m^* \lambda}{\kappa^3} \sinh \kappa \Delta\tau + \frac{q_m^* \lambda \Delta\tau}{\kappa^2} \cosh \kappa \Delta\tau, \end{aligned} \quad (4.27)$$

where we keep the notations from eq. (4.18):  $\kappa = \sqrt{-|q_m|^2 - \lambda^2}$ . For the AL method, the derivative of the transfer matrix acquires the following form:

$$\mathbf{T}'_m^{(\text{AL})} = \frac{1}{\sqrt{1 + \Delta\tau^2 q_m^2}} \begin{pmatrix} -i\Delta\tau e^{-i\lambda\Delta\tau} & 0 \\ 0 & i\Delta\tau e^{i\lambda\Delta\tau} \end{pmatrix}. \quad (4.28)$$

For the modified BO method we have

$$\mathbf{T}'_m^{(\text{BOmod})} = 2i\tau \sin |q_m \Delta\tau| \begin{pmatrix} 0 & e^{i(\theta_{q_m} + 2\lambda\tau)} \\ e^{-i(\theta_{q_m} + 2\lambda\tau)} & 0 \end{pmatrix}, \quad (4.29)$$

where  $\theta_{q_m}$  is an argument of  $q_m$ . For the modified AL matrix we have:

$$\mathbf{T}'_m^{(\text{ALmod})} = \frac{2i\tau \Delta\tau}{\sqrt{1 + \Delta\tau^2 |q_m|^2}} \begin{pmatrix} 0 & q_m e^{2i\lambda\tau} \\ q_m^* e^{-2i\lambda\tau} & 0 \end{pmatrix}. \quad (4.30)$$

The CN method transfer matrix derivative has the form [151]:

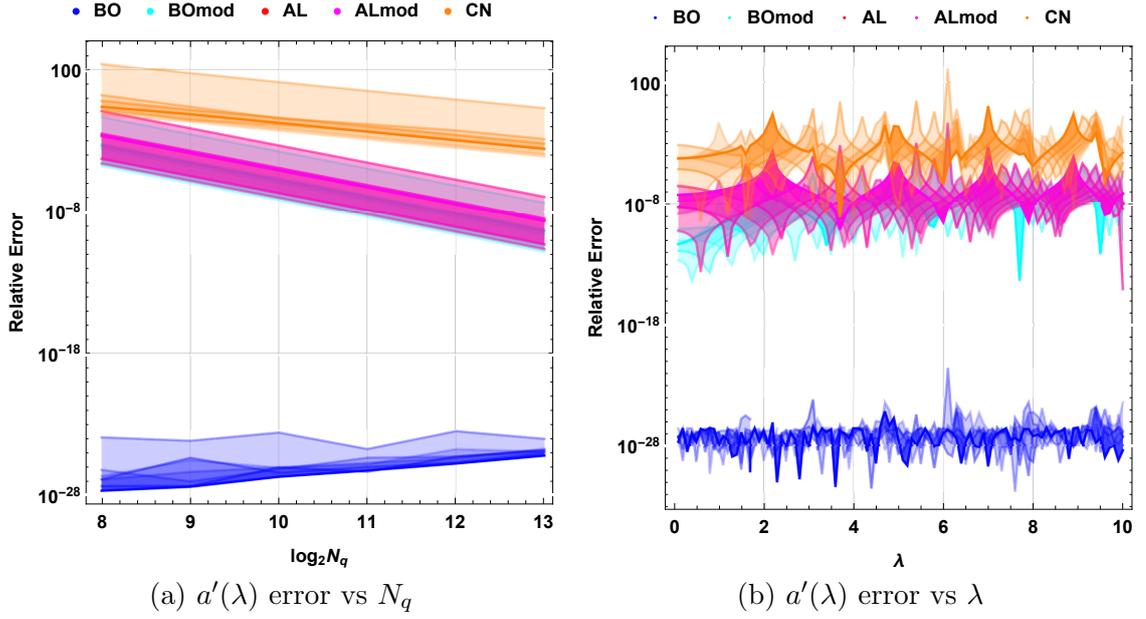
$$\mathbf{T}'_m^{(\text{CN})} = \frac{1}{2} \mathbf{P}_{m+1} (\mathbf{I} - \frac{\Delta\tau}{2} \mathbf{P}_{m+1})^{-2} (\mathbf{I} + \frac{\Delta\tau}{2} \mathbf{P}_m) + \frac{1}{2} (\mathbf{I} - \frac{\Delta\tau}{2} \mathbf{P}_{m+1})^{-1} \mathbf{P}_m, \quad (4.31)$$

where we again keep the notations from (4.24). In order to find  $a'(\lambda)$ , an additional step must be taken at the end of the algorithm. For the ordinary ZSP (3.5), we have the following expression for the derivative of  $a(\lambda)$  that involves the elements of the Jost solution and its derivatives:  $a'(\lambda) = (\psi'_1(L, \lambda) + iL\psi_1(L, \lambda))e^{i\lambda L}$ ; in the case of the envelope ZSP (3.27), we arrive at:  $a'(\lambda) = \chi'_1(L, \lambda)$ .

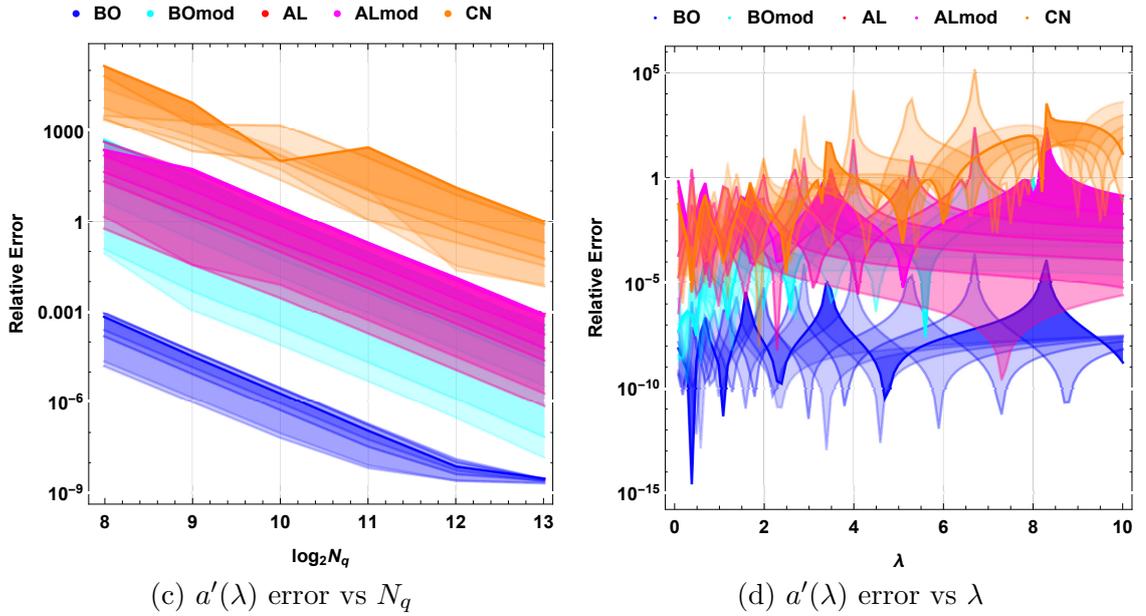
To check the accuracy of the derivatives evaluated via different algorithms, we resort to similar mechanisms as described above. We primarily use the scattering coefficient  $a(\xi)$  for the rectangular (4.8) and over-soliton (4.2) potentials. In figs. 4.5, we present the dependence of the relative error for the derivative  $a'(\lambda)$  on  $\lambda$  and on the number of subintervals  $N_q$  for the different methods described above. Our computations confirm that the BO method is again the most accurate one, while the accuracy of the CN is the lowest among all methods studied.

### 4.3 Iterative methods for discrete eigenvalues

Apart from the continuous spectral data computation, the full mapping to NF domain requires identification of all solitonic modes, i.e. discrete spectrum components. The requirement to ensure that there is no missed eigenvalue is the most challenging for the numerical studies of random signals. However, in communications, it is usually sufficient



$$q_{\text{rec}}, L = 1, A \in [2, 5], \Delta A = 0.5$$



$$q_{\text{over}}, L = 20, A \in [2.25, 5.25], \Delta A = 0.5$$

Figure 4.5: Relative error for  $a'(\lambda)$  computation as a function of the spectral parameter  $\lambda$  and the number of discretisation points  $N_q$  for different NFT methods and signal amplitude  $A$ .

to reveal the existence of the bound states from the pre-defined and known set, i.e. a communication alphabet [50, 54].

For the eigenvalues computation, the most popular option among the NFT related works is to apply some iterative scheme to identify the complex zero(s) of a function (namely,  $a(\lambda)$  from eq. (3.17)) in the case considered [151]. In [19, 141, 151], the dependence of the NFT method performance on a particular iterative scheme usage was somewhat overlooked and an arbitrarily chosen method was usually employed without a

particular motivation for the choice.

Traditionally, the most common iterative approaches are the secant and Newton-Raphson (NR) methods. Assuming an initial guess for the location of the zero  $x_0$  of some function, say  $f(x)$ , the consecutive iteration scheme for the NR method is given by:

$$x_{k+1}^{(\text{NR})} = x_k - \frac{f_k}{f'_k}, \quad (4.32)$$

where  $f_k := f(x_k)$ . This method has a quadratic convergence rate [60] (numerical method is said to have convergence rate  $p$ , if  $|x_{k+1} - x| \leq C|x_k - x|^p$ , here and below the orders of convergence are given under the assumption that all roots are simple). The main disadvantage of the NR method is the necessity to know the value of the function derivative at each iteration step. For the purpose of brevity, we introduce a shorthand notation for the so-called divided differences:

$$f[x_1, x_2] = \frac{f(x_2) - f(x_1)}{x_2 - x_1}, \quad (4.33)$$

which can be recursively generalized for an arbitrary number of arguments:

$$f[x_k, \dots, x_{k+m}] = \frac{f[x_{k+1}, \dots, x_{k+m}] - f[x_k, \dots, x_{k+m-1}]}{x_{k+m} - x_k}. \quad (4.34)$$

In the secant method, the expression for the derivative in each iteration is swapped over to the divided difference, leading to:

$$x_{k+1}^{(\text{secant})} = x_k - \frac{f_k}{f[x_{k-1}, x_k]} = x_k - f_k \frac{x_k - x_{k-1}}{f_k - f_{k-1}}. \quad (4.35)$$

The convergence rate of the secant method is approximately 1.618, so it is worse than that for the NR method, but a single step computation using the secant method can be faster since it does not require computing the derivatives. Sidi [129] generalised the idea of the derivative approximation: the function derivative  $f'_k$  can be replaced by the derivative of a fitting polynomial  $p(x)$  of degree  $j$ :

$$p'_j(x) = f[x_{k-1}, x_k] + \sum_{i=2}^j f[x_{k-i}, \dots, x_k] \prod_{l=1}^{i-1} (x_k - x_{k-l}). \quad (4.36)$$

The next iteration is given by

$$x_{k+1}^{(\text{Sidi})} = x_k - \frac{f_k}{p'_j}. \quad (4.37)$$

In our study, we use a cubic polynomial approximation in (4.36), i.e.  $j = 3$ . For this particular case, the convergence rate of the method is  $\approx 1.93$  [129].

Steffensen's method [62] uses the following iterative formula:

$$x_{k+1}^{(\text{Steffensen})} = x_k - \frac{f_k^2}{f(x_k + f_k) - f_k}. \quad (4.38)$$

It allows us to reach the convergence rate 2, same as that for the NR method.

The Muller method [104] has an advantage in that it allows us to find complex roots from a real initial guess. Defining the auxiliary quantities

$$\begin{aligned} w &= f[x_{k-1}, x_k] + f[x_{k-2}, x_k] - f[x_{k-2}, x_{k-1}], & g &= f[x_{k-2}, x_{k-1}, x_k], \\ d &= \max \left[ w - \sqrt{w^2 - 4f_k g}, w + \sqrt{w^2 - 4f_k g} \right], \end{aligned} \quad (4.39)$$

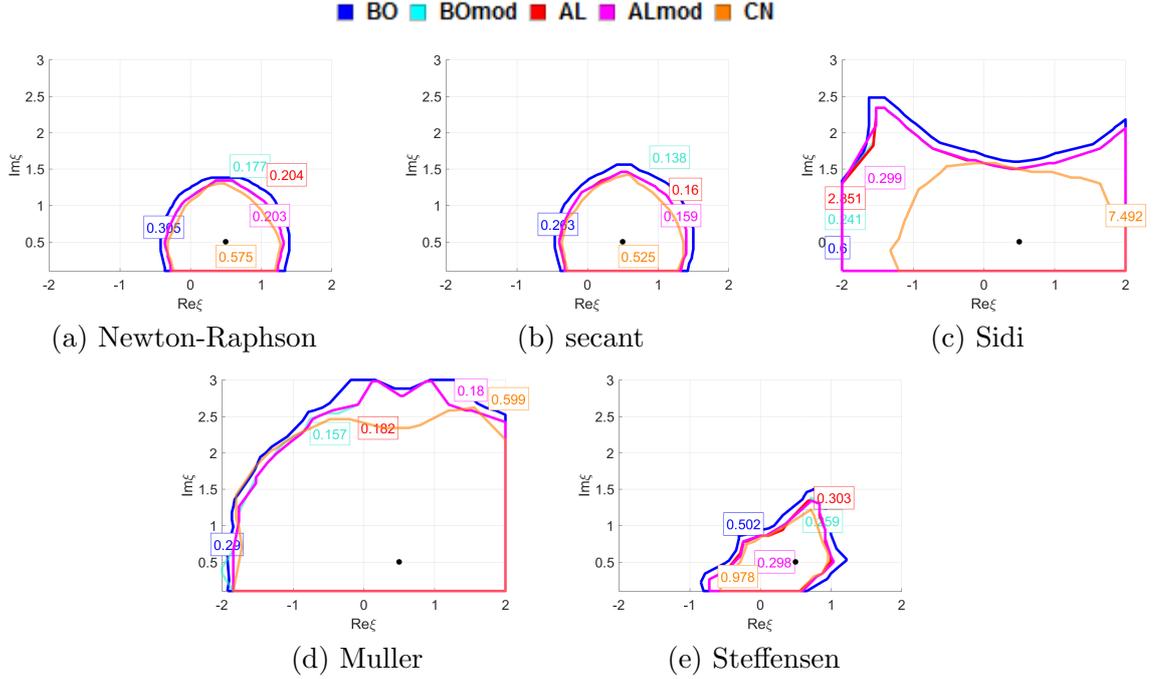


Figure 4.6: The regions of initial assumptions for the zero approximation in which the iterative algorithms reach correct zero (marked as black point) with less than 1% relative error. The computation is performed for  $q_{\text{sol}}$ ,  $N_q = 2^{10}$ ,  $L = 30$ . The digits in the labels identify the average computational runtime. Perfect vertical and horizontal edges of some basins mean that regions of convergence extend out of the simulation ranges.

(the maximum is determined by comparing the absolute values), the iteration step of the Muller method is given by

$$x_{k+1}^{(\text{Muller})} = x_k - 2 \cdot f_k/d. \quad (4.40)$$

The order of convergence for this method is approximately 1.84 [60], which is better than that for the secant method.

All iterative algorithms applied for the computation of the eigenvalues require good initial guess. In order to understand how the choice of the initial value influences the result of the eigenvalues search, I performed a study of the convergence of all iterative schemes in dependence on the initial guess point value: fig. 4.6 contains the results referring to the frequency-shifted soliton potential, eq. (4.13), and fig. 4.7 depicts the results for the rectangular potential, eq. (4.7), with  $A = \pi/2$  and  $L = 1$ , where only a single eigenvalue is present. On these plots, we show the border of the regions in the complex plane of spectral parameter  $\lambda$  (marked with the closed lines of a different colour), where the relative error of zero location is less than 0.01. The iterative algorithm runs until it reaches the pre-set precision in the difference between the function values for consequent iterations or until it exceeds the pre-set number of iterations (these pre-sets were, correspondingly,  $10^{-10}$  and  $10^3$ ). Each line on the plots is labelled the corresponding average runtime that the computation of the eigenvalue takes when the initial guess point is positioned inside the respective regions.

The result of our analysis of the aforementioned five iterative methods combined with different NFT algorithms can be summarized as follows: the NR and secant methods, used in [141, 151], have smaller regions of convergence than those of the Sidi and Muller methods (the Muller method shows the largest convergence basin for all example profiles used), whereas the runtimes for all four approaches are similar. In the case of the Sidi algorithm, we observe a better convergence, but with significantly higher runtime, especially when

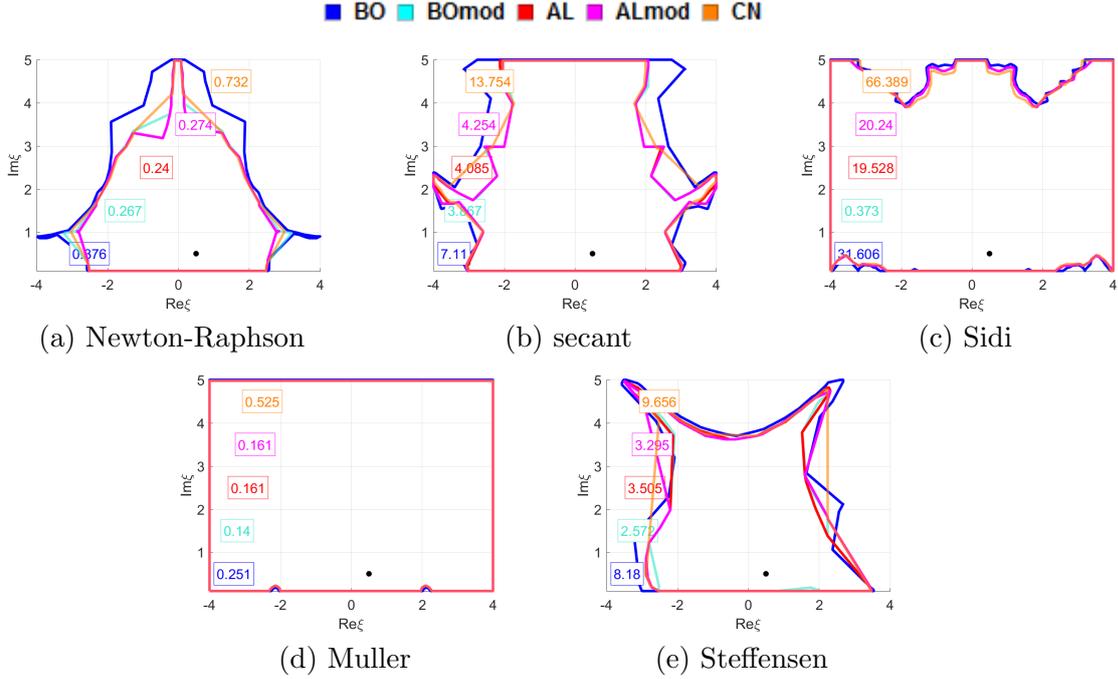


Figure 4.7: The regions of initial assumptions for the zero approximation in which the iterative algorithms reach correct zero (marked as black point) with less than 1% relative error. The computation is performed for  $q_{rec}$ ,  $N_q = 2^{10}$ ,  $L = 1$ ,  $A = \pi/2$ . The digits in the labels identify the average computational runtime. Perfect vertical and horizontal edges of some basins mean that regions of convergence extend out of the simulation ranges.

the iterative method is coupled with the AL and CN algorithms. The Steffensen's iterative method shows the even smaller region of convergence, which is worse than that for the NR and secant methods.

The drawback of iterative methods in the application to real transmission problems (when the position of solitonic eigenvalues is usually not known a priori) is that we cannot predict the computational runtime since we cannot estimate the number of iterations required to find the zero point with satisfactory accuracy. Moreover, the methods can fail to converge at all, so that some additional precautions have to be taken. The previous works devoted to the eigenvalue search algorithms assumed iterative schemes for estimating discrete eigenvalues. The main results referred to particular features of the NFT method but not necessarily to a particular iterative scheme. Authors of [151] studied the one-soliton and multisoliton cases for the AL, Euler and RK methods (also using the CN method for some cases, but evidently revealing its weaknesses). They also found out that all studied root-search techniques resulted in similar accuracy regimes for the NF data computed.

According to [19], the RK algorithm can converge faster than the BO, but the authors there used the grid search for the location of eigenvalues, and this resulted in a relatively high overall runtime. It was found that in the case of rectangular potential, the RK method failed in the correct computation for the total number of zeros (the authors used the total phase increment along the  $\Re\lambda$  axis for this purpose).

Our analysis of the iterative algorithms reveals that they are not sufficiently stable and manageable for the eigenvalue computations in realistic applications. This fact motivated the search for principally different methods and options applicable to the location of eigenvalues. Meanwhile, we understand that the region (in the complex  $\lambda$  plane) of the possible localisation of the eigenvalues should be another input of the algorithm, implemented for

their search. We refer to this region as region of interest (ROI) below.

In the case of several eigenvalues, embedded to the profile, the iterative search must be adjusted, because it provides a search of the zeros one by one. The concept is to eliminated already found zeros of the function  $a(\lambda)$  from the set of its zeros. One of the possible ways of doing this is provided in algorithm 1. Notice that redefinition of the function of interest in the iterative search may introduce additional numerical issues in the proximity of the located eigenvalue. The multieigenvalue case will be addressed more in sections 5.2 and 5.3.

---

**Algorithm 1** Adaptation of an iterative algorithm to the multisolitonic case

---

**Input:** Define function  $f(x)$ , make initial guess  $x_0$ ,  
 expected number of zeros  $N_s$ , ROI.  
**Step 0:** Initialize an empty output array  $x_{\text{out}}$ .  
**Step 1:** Define a current function  $f_c(x) = f(x)$   
 and a current guess  $x_c = x_0$ .  
**Step 2:** Launch the iterative algorithm for  $f_c(x)$ ,  $x_c$   
 and limited number of iteration steps.  
**Step 3:** Check if zero  $x_i$  was successfully located in the previous step.  
 If yes, go to step 4.  
 If no, go to step 7.  
**Step 4:** Check if the located zero is inside the ROI.  
 If yes, concatenate the located zero to the output array  $x_{\text{out}} = [x_{\text{out}}, x_i]$   
 and go to step 5.  
 If no, go to step 7.  
**Step 5:** Check if all expected zeros are located using  $\text{size}(x_{\text{out}}) == N_s$ .  
 If yes, go to Output.  
 If no, go to step 6.  
**Step 6:** Redefine the function to eliminate located zero  
 $f_c(x) = f_c(x)/(x - x_i)$ .  
**Step 7:** Choose an initial guess number  $x_c$  randomly from inside the ROI  
 and go to step 2.  
**Output:** Estimate the array of zeros  $x_{\text{out}}$ .

---

## 4.4 Fourier collocation method

In this subsection, I provide the previously not published results in the so-called Fourier collocation (FC) method performance analysis and the derivation of the updated version. Also, the subsection contains a comparison between two of the and discussion of the applicability.

The procedure of finding the discrete NF eigenvalues can be reformulated in terms of finding the eigenvalues of the eigenproblem corresponding to the discretised version of ZSP, this approach is known as FC or spectral method [144, 151]. To formulate the ZSP as an eigenproblem, one should rewrite eq. (3.5) as in eqs. (3.1), (3.4):

$$-i \begin{pmatrix} -\frac{d}{d\tau} & q(\tau) \\ q(\tau)^* & \frac{d}{d\tau} \end{pmatrix} \Psi(\tau, \lambda) = \lambda \Psi(\tau, \lambda). \quad (4.41)$$

Now one decomposes the sought solution  $\Psi$  and the potential  $q$  (defined on the finite interval  $\tau = -L \dots L$ ) into the Fourier series of length  $2n + 1$ , then the discrete version of (4.41) in the Fourier space transforms into the eigenvalue problem for a block matrix,

where each block corresponds to each respective element of matrix in the l.h.s. of (4.41). The derivative  $-id/d\tau$  results in the block  $\mathbf{B1}$ :

$$\mathbf{B1} = k_0 \text{diag}(-N..N), \quad (4.42)$$

where  $k_0$  is the resolution frequency of Fourier series:  $k_0 = 2\pi/2L$ ,  $\text{diag}(\cdot)$  is a diagonal matrix with the parenthesis's content along the diagonal, and  $n$  in the half-number of Fourier modes. The potential  $-iq(\tau)$  is presented as

$$-iq(\tau) = \sum_{j=-n}^n c_j e^{ijk\tau}, \quad (4.43)$$

and the coefficients of this decomposition,  $c_j$ , form Toeplitz matrix  $\mathbf{B2}$  of the dimension  $(2n+1) \times (2n+1)$ , cf. eq (5):

$$\mathbf{B2} = \text{Toeplitz}(c_0, \dots, c_n, 0, \dots, 0; c_0, \dots, c_{-n}, 0, \dots, 0). \quad (4.44)$$

where the expression in the parenthesis containing  $n+2$  elements is the first row (column) of the Toeplitz matrix defining its full structure. Now we deal with the eigenproblem for the resulting block matrix:

$$\mathbf{B} = \begin{pmatrix} -\mathbf{B1} & \mathbf{B2} \\ \mathbf{B2}^+ & \mathbf{B1} \end{pmatrix}. \quad (4.45)$$

The algorithms of eigenvalues search in MATLAB are based on the LU factorisation [116]. The complexity of such an algorithm for the non-Hermitian matrix grows with the matrix's rank as  $O((\text{rank } \mathbf{B})^3)$  (perhaps, in practice this approach is extremely fast). In addition, the algorithm returns by default rank  $\mathbf{B}$  eigenvalues. Some resulting eigenvalues are due to be discarded because they are located in lower half-plane. After that additional analysis must be performed to distinguish the true solitonic eigenvalues (with the positive imaginary part) from the continuous NF modes (corresponding to purely real eigenvalues), when the real eigenvalues have acquired a spurious small imaginary part as a result of the finite numerical precision. In practical applications, the eigenvalues typically lie sufficiently far from the real axis. Thus, in order to remove spurious eigenvalues with small positive imaginary parts some additional condition on  $\Im\lambda$  can be used to guarantee that the imaginary part is above some predefined threshold value. It is also known that for a single-lobe potentials (like an over-soliton or a rectangle pulse), the eigenvalues are pure imaginary [60]. This condition can also be used to get rid of spurious eigenvalues albeit in the true situation, like the NFT processing of realistic profiles in the optical transmission tasks when the profiles are usually quite involved and can contain some noisy component [117], only the threshold for the imaginary part of eigenvalues can be efficiently used.

In this thesis, for the first time, I present the modification and improvement of the method based on the eigenvalues search inside the Fourier space. Since eigenvalues are found from the determinant and corresponding characteristic equation, we start with the decomposition of the block-matrix determinant [132]:

$$\det \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \det \mathbf{A} \cdot \det(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}). \quad (4.46)$$

In order to apply this relation to equation defining the desired eigenvalues,

$$0 = \det(\mathbf{B} - \mathbf{I}\lambda) = \det \begin{pmatrix} -\mathbf{B1} - \mathbf{I}\lambda & \mathbf{B2} \\ \mathbf{B2}^+ & \mathbf{B1} - \mathbf{I}\lambda \end{pmatrix}, \quad (4.47)$$

it is more convenient to rewrite matrix in the determinant using the following transformations:

$$\begin{aligned} 0 &\sim \det \begin{pmatrix} \mathbf{B2} & -\mathbf{B1} - \mathbf{I}\lambda \\ \mathbf{B1} - \mathbf{I}\lambda & \mathbf{B2}^+ \end{pmatrix} = \\ &\sim \det \mathbf{B2} \cdot \det(\mathbf{B2}^+ - (\mathbf{B1} - \mathbf{I}\lambda)\mathbf{B2}^{-1}(-\mathbf{B1} - \mathbf{I}\lambda)). \end{aligned} \quad (4.48)$$

These transformations influence only the overall sign of the expression and do not change characteristics equation. After using (4.46), one can obtain:

$$0 = \det \mathbf{B2} \cdot \det(\lambda^2 \mathbf{B2}^{-1} + \lambda(\mathbf{B2}^{-1}\mathbf{B1} - \mathbf{B1B2}^{-1}) - \mathbf{B2}^+ - \mathbf{B1B2}^{-1}\mathbf{B1}). \quad (4.49)$$

The first determinant does not depend on  $\lambda$  and we can disregard it. finding the second determinant reduces to the solution of the so-called quadratic eigenproblem:

$$(\mathbf{M}\lambda^2 + \mathbf{P}\lambda + \mathbf{Q})\Phi = 0. \quad (4.50)$$

This problem can be readily solved numerically using the embedded `MATLAB` function `polyeig`. In application to current problem, the matrices in quadratic eigenvalues problem (4.50) are defined as:

$$\mathbf{M} = \mathbf{B2}^{-1}, \quad \mathbf{P} = \mathbf{B2}^{-1}\mathbf{B1} - \mathbf{B1B2}^{-1}, \quad \mathbf{Q} = -\mathbf{B2}^+ - \mathbf{B1B2}^{-1}\mathbf{B1}. \quad (4.51)$$

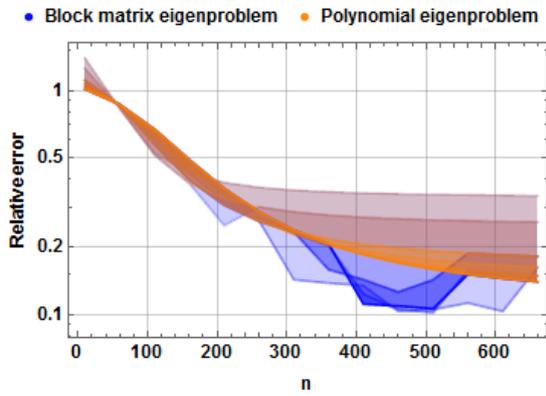
The problem can be simplified further in the case when  $\mathbf{B2}^{-1}$  and  $\mathbf{B1}$  commute. Note that after the simplification described above, the computations within our new method become faster as the rank of the resulting matrix decreases by the factor of 2. At the same time, as before, we still need to impose the additional conditions on the particular values of the eigenvalues obtained to extract exactly the eigenvalues that describe the solitons.

I performed series of tests to check how the upgraded FC method works against the previous version of the FC method (that assumed a straight digitalisation procedure), using the three model potentials from subsection 4.1. In the left panes of fig. 4.8, the relative error of the eigenvalue found versus the number of points used for the discretization of the example profiles. The right column of panes in the same figure gives the computation time  $t_{\text{run}}$  normalised to the number of samples  $n$  required to reach the result from the left panes.

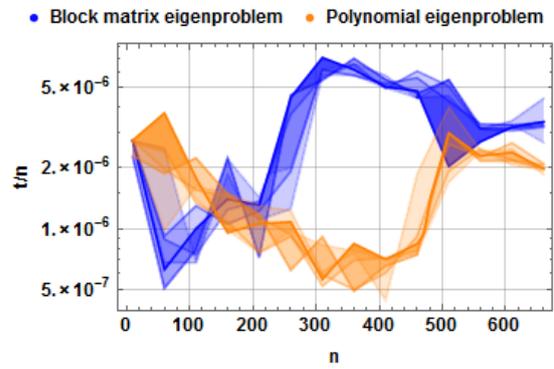
The test shows that the simplified FC polynomial eigenvalues method requires a smaller runtime needed to reach the result while demonstrating a similar accuracy as the straight diagonalisation. For the eigenvalues search employing single-lobe potential, the eigenvalues can be located easily and with high accuracy. However, for the solitonic potential with a phase factor, fig. 4.8 (e), (f), the new FC method does not provide the sufficient accuracy for the eigenvalue, but so far we have now found the explanation for this behaviour. At the same time we mention that, typically, the family of FC methods requires a considerably higher number of operations to get the result with the acceptable quantity as compared to the transfer-matrix methods described in subsection 4.2, such that further, we concentrate on the properties for the latter. It is known that Fourier collocation method has the so-called 'spectral accuracy', which can cause the presented results.

## 4.5 Discrete spectrum methods discussion

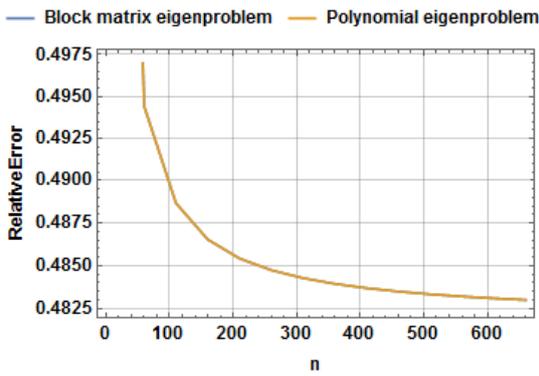
In the community of numerical methods for NFT and transmission based on NFT, the discrete spectrum attracts less attention than continuous spectral data. That happens, on one hand, because the discrete spectrum provides poorer opportunities to modulate the information and control the signal parameters. On the other hand, the very truncation of the potential on the finite time support contradicts with the idea of a solitonic wave,



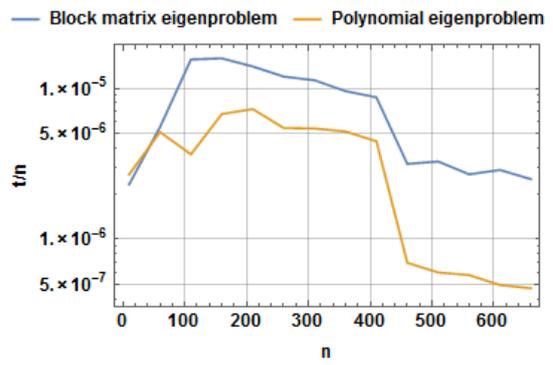
(a) eigenvalue's error for  $q_{\text{over}}, L = 20$



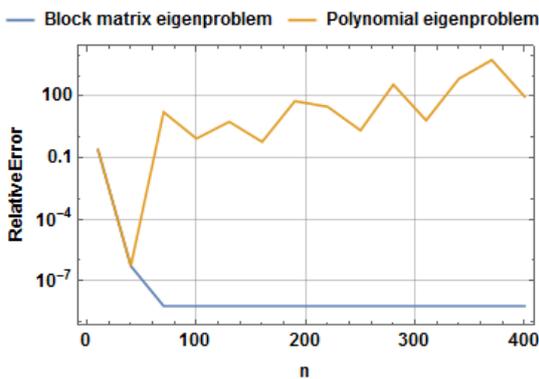
(b) normalised runtime for  $q_{\text{over}}, L = 20$



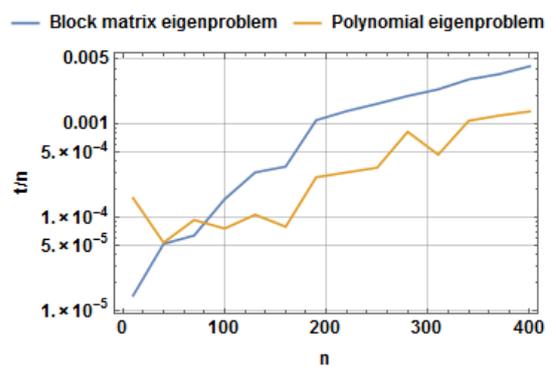
(c) eigenvalue's error for  $q_{\text{rec}}, L = 1, A = \pi/2$



(d) normalised runtime for  $q_{\text{rec}}, L = 1, A = \pi/2$



(e) eigenvalue's error for  $q_{\text{sol}}, L = 20$



(f) normalised for  $q_{\text{sol}}, L = 20$

Figure 4.8: Relative error and normalised computational time of discrete eigenvalue computed using two types of FC method in dependence on number of points  $n$ .

and soliton tails decay is not sufficient for many soliton profiles. More explicitly, in order to get localisation of solitonic pulse, its amplitude should be rather high. It leads to an increase in the numerical error and higher signal peak-to-average power ratio (PAPR). Besides, the problem of designing the desired waveform by adjusting the discrete spectrum components is still open, and the computation of even 10-soliton profile requires to move from double-precision to higher accuracy. From the point of direct problem, there is no way to ensure that all solitonic modes of the processed signal are revealed, there are always can be some low amplitude soliton, not recognisable because of the truncation or finite sampling resolution.

All these facts make the usage of solitonic modes for general signal processing rather impractical. The communication application provides some apriori information: we usually start from constructing the signal with the given alphabet of available solitonic modes, where we can either use all of given eigenvalues, playing with the solitonic phases, or some subset of them, playing with number and location of chosen eigenvalues. The latter approach allows more degrees of freedom (for  $N_s$  eigenvalues, one can either modulate  $N_s$  phases, or  $2^{N_s}$  combinations of eigenvalues), however, the signal power becomes input-dependent.

Another point is the relation between asymptotical complexity and actual numerical methods runtime. The FC method is extremely fast, despite its complexity is  $O(n^3)$ . However, there is no clear relation, how many Fourier components  $n$  should be used for the signal. For any practical realisation of the transmission system, this number is manually adjusted to allow processing of the discrete eigenvalues from the chosen alphabet. The iterative methods are also rather practical if we preliminary have the information about the region of the location of the eigenvalues. Finally, all existent methods are useful for communication applications, due to the apriori knowledge. In general situation, both approaches require noticeable efforts in adjusting the parameters of algorithms and lead to rather unpredictable complexity, runtime and outcome. This motivation moves us to develop a more universal approach to the location of the eigenvalues, described and discussed in the following chapter.

## Chapter 5

# Contour integrals within inverse scattering problem

In this chapter, I present one of the main results of my PhD research. It lies in the application of contour integrals (in NF domain) to the detection of scattering data, or to information modulation. This method shows good performance in terms of computational complexity, almost independent on the number of eigenvalues, numerical accuracy, effective dealing with noisy systems or systems with closely located eigenvalues.

The first introduction of this method has been done in the large review paper of numerical methods for NFT [J2], and the properties of the noise tolerance and degenerative eigenvalues detection were presented in [J3]. Also, I have presented this approach and its advantages on the Institute of Mathematics and its Applications (IMA) Conference for Nonlinearity and Coherent Structures in 2017 [C8]. Besides, this chapter contains unpublished results about the design of the contour integrals based communication system.

### 5.1 Contour integrals for zero location

#### 5.1.1 Mathematical details

Initially, the contour integrals approach was presented as a numerical tool for localisation of zeros of a nonlinear function in the complex plane by Delves and Lyness (DL) in [30], together with two complementary papers for numerical differentiation of an analytical function [93, 95], and one complimentary paper on numerical contour integration [94]. They utilized the symmetrical polynomial dependence between the sequence of integrals  $\{s_p\}_{p=1}^{N_s}$ , defined as

$$s_p = \frac{1}{2\pi i} \int_C x^p \frac{f'(x)}{f(x)} dx, \quad (5.1)$$

(for manually chosen integration contour  $C$ ), and zeros of the function  $f(x)$ , which lie inside the given contour  $C$ , so  $s_p = \sum_{j=1}^{N_s} x_j^p$ , where  $x_j$  are abovestated zeros. Note that there is no need to know the exact number of zeros inside the contour, since zero-order integral returns this values as  $s_0 = \frac{1}{2\pi i} \int_C x \frac{f'(x)}{f(x)} dx = N_s$ . This number takes into account zeros multiplicity (if there is any degenerate zeros). This particular property of the analytic function is well known in the NFT community and widely used for the number of eigenvalues estimation, see, e.g., [19].

The evaluation of  $\{s_p\}$  allows deriving the so-called Newton's identities  $\{\kappa_p\}_{p=1}^{N_s}$ :

$$\begin{aligned}\kappa_1 &= -\sum_{j=1}^{N_s} x_j, \\ N_s \kappa_2 &= x_1 x_2 + x_2 x_3 + \dots + x_{N_s-1} x_{N_s}, \\ &\dots \\ \kappa_{N_s} &= (-1)^{N_s} x_1 x_2 \dots x_{N_s},\end{aligned}\tag{5.2}$$

which are linked to contour integrals  $s_p$  as

$$\begin{aligned}s_1 + \kappa_1 &= 0, \\ s_2 + s_1 \kappa_1 + 2\kappa_2 &= 0, \\ &\dots \\ s_{N_s} + s_{N_s-1} \kappa_1 + \dots + s_1 \kappa_{N_s-1} + N_s \kappa_{N_s} &= 0.\end{aligned}\tag{5.3}$$

This system of equations can be solved recurrently using the values of  $\kappa_p$  obtained in the previous iteration:

$$\kappa_p = \frac{1}{p} \left( s_p + \sum_{j=1}^{p-1} s_j \kappa_{p-j} \right), \quad p = 1 \dots N_s.\tag{5.4}$$

The same  $\kappa_p$ -s are (up to a sign) the Vieta's formulae for the following polynomial:

$$M(z) = z^{N_s} + \kappa_1 z^{N_s-1} + \kappa_2 z^{N_s-2} + \dots + \kappa_{N_s-1} z + \kappa_{N_s}.\tag{5.5}$$

The polynomial  $M(z)$  has exactly the same zeros as the initial function  $f(x)$ . Therefore, using any polynomial root-finding technique, the desired set of zeros inside the contour can be computed easily.

With regards to the proper choice of the integration contour, it was reported in [30] that the numerical error of contour integrals approach is inversely proportional to the distance between the contour and the nearest zero [30]. That means that choice is based on the trade-off between higher accuracy of zero evaluation and higher runtime (more points of the function  $f(x)$  evaluation) for larger contour.

Since the solitonic eigenvalues are defined as zeros of auxiliary scattering function  $a(\lambda)$ , eq. (3.17), I successfully applied DL approach to computation of discrete NF spectrum. There are two reasons, why this method is effective. First, it naturally includes the region of zeros location. That is convenient because the solitonic eigenvalues have certain limitations in spreading in the complex plane: they are locating in the upper half-plane and cannot have a too large imaginary part, because it is associated with the solitonic mode energy (in particular, for pure soliton, energy is given by  $2A = 4\Im\lambda_{\text{sol}}$ ), see subsection 3.1.3. Moreover, in communication application, it is unlikely to employ the solitonic modes with large real part as well, because it increases the signal bandwidth and may lead to escaping of the mode from the dedicated time window. So, this approach acts as effective filtering of the spurious eigenmodes, which are shown to be arising during the signal propagation, because of deviations of the channel model from the purely integrable [31].

Another reason, which makes DL approach a promising tool, is its independence on the number and spreading of the eigenvalues within the contour's interior. The algorithm requires only one round of scattering data (values of  $a'(\lambda)/a(\lambda)$ ) computation, for the values of NF frequency along the contour. Then all evaluation steps are based on the computed values, however many eigenvalues lie inside the contour. Also, this approach easily resolves the case of degenerative eigenvalue, because zeros of the function of interest

are returned simultaneously together with their multiplicity. One can also intuitively expect more stability of the overall algorithm, because the scattering data are computed for the points of  $\lambda$ -plane far from the singularities, where the numerical stability of NFT algorithm may be an issue.

There are two main competitors of the contour integrals approach as a method for discrete NF spectrum computation: spectral or Fourier Collocation method, see section 4.4 and the group of various iterative methods, section 4.3. In terms of the computational complexity, I expect to get an advantage in realistic applications. The iterative algorithm, implemented alone, requires  $O(N_q \cdot N_s N_{\text{iter}})$  operations, where  $N_{\text{iter}}$  is the number of steps to reach the desired zero with a given accuracy. But when we have an insufficiently accurate initial guess, the number of required iterative steps  $N_{\text{iter}}$  can be dramatically large, going to infinity when the algorithm cannot converge.

More recently, Kravanja et al in [75] presented an improved version of the DL approach. Their algorithm relied on a recursive construction of the so-called formal orthogonal polynomials, which have the roots that coincide with the zeros of our function  $f(x)$ . I have analysed the accuracy of this improved method for eigenvalues computation too, but we do not present these results here on separate plots because our study did not reveal any noticeable difference between results obtained from this newer method compared to those of the ordinary DL method (at least, for the set of our test profiles). However, for some real-world applications, where signals are not smooth and often significantly corrupted by noise [135], the approach proposed in [75] might demonstrate better performance.

### 5.1.2 Numerical accuracy of the contour integrals for eigenvalues location

The DL approach allows at least two variations in applications to the NFT problems. Note that the nonlinear function of interest  $f(x)$  in eq. (5.1) is involved together with its derivative, so, to use it for eigenvalues, one needs to integrate over the contour a logarithmic derivative  $a'(\lambda)/a(\lambda)$ . Since the computation of the auxiliary scattering function is the computationally consuming task here, it should be carefully considered. One can use the techniques from subsection 4.2.1 to evaluate the proper values of  $a'(\lambda)$  for the points along contour. Alternatively, one can approximate the derivative with the discrete difference, substituting in the integrand:

$$x^p \frac{a'(x)}{a(x)} dx \rightarrow x_k^p \frac{a'(x_k) \Delta x}{a(x_k)} \approx x_k^p \frac{a(x_k) - a(x_{k-1})}{a(x_k)} = x_k^p \left( 1 - \frac{a(x_{k-1})}{a(x_k)} \right). \quad (5.6)$$

I label this approximation method used for the search of solitonic eigenvalues with integrand approximated as aDL.

We analysed the dependence of contour integration algorithms' accuracy on the number of points along the contour and on the contour shape, fig. 5.1. We here assume that the contour envelopes the region where the eigenvalues can be located, ROI in other words. We study the behaviour of the aDL method with the approximation scheme represented in eq. (5.6) and compare it with the ordinary DL method's behaviour. As it was expected, the latter works more accurately due to a more accurate calculation of the derivative, see the blue and orange lines and compare them with red and green ones in fig. 5.1. We also observed that the particular contour shape was chosen in eq. (5.1) also influences the resulting accuracy of the eigenvalues found. In particular, we checked the behaviour of the methods using the rectangular contour in the upper half-plane of  $\lambda$ , fixing the contour borders along  $\Re\lambda$ -axis and  $\Im\lambda$ -axis. Another option that we tested was to define the ring sector in the  $\lambda$ -plane, fixing the borders for absolute value  $\rho$  and for argument  $\theta$  of  $\lambda$  written in polar representation as  $\lambda = \rho e^{i\theta}$ .

The integrals in all of our methods have been evaluated using the trapezoidal rule.



in just several iterative steps, ensuring a lower computational time.

## 5.2 Multisolitonic test of the discrete eigenvalues methods

Multisolitonic potentials are patently interesting from the perspective of practical applications, with subsets of multisolitonic eigenvalues being specifically proposed for optical communication purposes [50, 51, 54]. In this subsection, we investigate the performance of both iterative and contour integration methods for the case when the pulse decomposition involves several eigenvalues, and we need to retrieve their values. For our tests, we choose the over-soliton potential as in eq. (4.1) for the fixed amplitude  $A = 5$ . This signal has five eigenvalues in its discrete NF spectrum:

$$\lambda_k = (4.5 - k)i, \quad k = 0 \dots 4. \quad (5.7)$$

The full decomposition also contains the non-zero continuous spectral data, so that the situation considered in this section is quite general. To locate the eigenvalues, it is convenient to use the rectangular ROI for the contour integration methods because, as it was found from fig. 5.1, such a contour ensures greater accuracy in results. We choose the contour having the shape of a rectangle in the complex  $\lambda$ -plane with the dimensions:  $[-1, 1]$  along the real axis and  $[0.1, 5]$  along the imaginary axis. Such a rectangle encompasses all eigenvalues defined by eq. (5.7) for the over-soliton profile.

In our specific case, we started with the initial guess  $\lambda_0 = i$ , as in algorithm 1, followed by up to three attempts to locate the same zero over  $10^5$  iteration steps to reach each particular zero point.

The resultant error versus runtime diagram is represented as a bubble chart in fig. 5.2 for the different combinations of the transfer-matrix algorithms (subsection 4.2) and the root-finding methods (subsections 4.3 and 5.1). Here bubble sizes are inversely proportional to the runtime with the numbers therein indicating the number of zeros; respective colours identify the numerical algorithm chosen.

We see that, in accordance with the remarks made in the previous subsection, the iterative algorithms equipped with the elimination procedure may fail to estimate the entire cluster of zeros even when coupled with the most accurate BO method. At the same time, the contour integration methods, especially aDL, give higher error margin.

## 5.3 The contour integrals method paired with iterative search

Together with the application of the contour integrals approach solely, one can use their output to get a better initial guess for the iterative algorithm. It is motivated by the fact that both types of eigenvalue-finding approaches have some inherent disadvantages. The iterative algorithms solely can be unstable and no one can ensure that the iterations will eventually lead to the correct eigenvalue points. They also require some additional adaptation to the multisolitonic case to incorporate the elimination of previously found zeros, and the restriction on the search region has to be generally imposed. The pure contour integration algorithms do not allow us to reach high accuracy and take a comparatively long time.

To overcome these difficulties, I introduced a hybrid method, which allows taking advantage of the best from both approaches whilst simultaneously getting rid of their respective drawbacks. The main idea of our hybrid method is that we can use the result of contour integration as the initial guess that is then supplied to the consequential iterative algorithm (the same strategy was mentioned in [30] as a way of root-finding refining). This combination allows us to reach almost any accuracy up to the limitations imposed by

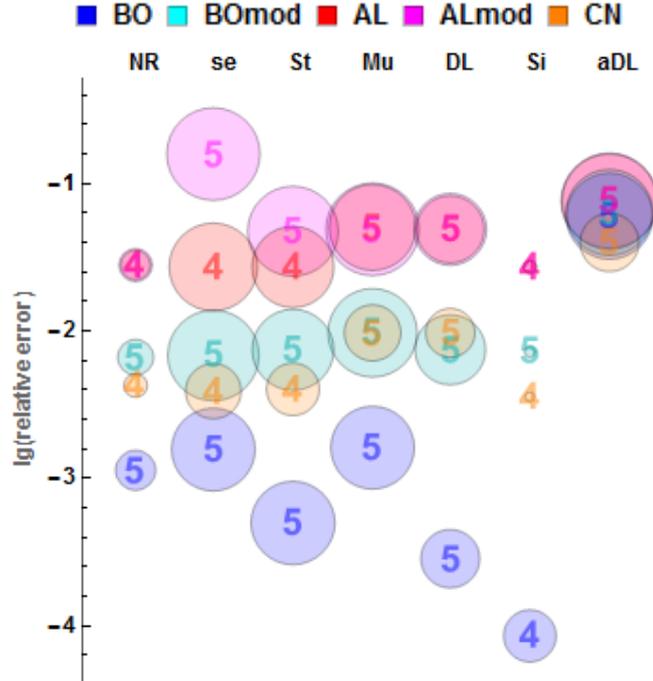


Figure 5.2: Relative error of the eigenvalue location via different root-finding iterative, NR, se(cant), St(effensen's), Mu(ller), Si(di), and contour integrals, DL, aDL, algorithms for over-soliton profile with 5 eigenvalues,  $q_{\text{over}}$ ,  $A = 5.25$ ,  $L = 20$ . The digit in each bubble means the number of the located zeros (out of 5). The bubble size shows the inverse runtime (a smaller bubble means a longer run and vice versa), runtime changes from 0.025 s to 4192 s.

the NFT computation method itself, see section 4.2. The hybrid method presented here guarantees locations for all eigenvalues as opposed to iterative methods that often found most eigenvalues but not necessarily all. The description of the consecutive steps for the hybrid algorithm is as follows.

- First, find the location of the approximate zeros' using one of the contour integration method drawing a large enough contour. Two key remarks on this step here:
  - the integration result does not require to be really accurate, so the computation time can be reduced;
  - it allows us to find a good approximation for all zeros that lie inside the contour of interest;
- Second, we apply a particular iterative method to find more precise location for each eigenvalue, using the results of the previous step as guess points (also employing some other data that can be obtained by the contour integration, i.e. the multiplicity).

To test the hybrid algorithm, I use the over-soliton potential with five solitonic modes  $q(\tau) = 5.25 \text{sech}\tau$  as a test profile, and plot the error of the eigenvalues computed in fig. 5.3, using different combination of the transfer-matrix methods for ZSP (see section 4.2), iterative methods and contour integrals in the DL and aDL forms.

In general, runtime reduces comparably with contour integrals approach, but it is still larger than the runtime of some iterative algorithms (see table 5.2 for the runtimes of BO method used as an example). The largest runtime is demonstrated by the Sidi and Steffensen's methods, but the latter show a surprisingly high accuracy when coupled with

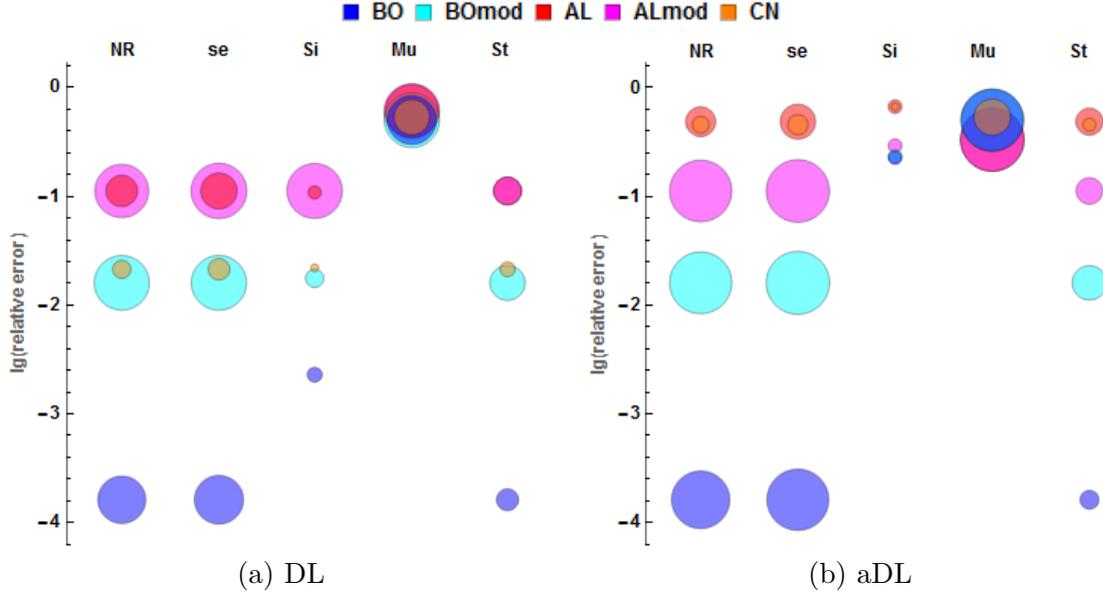


Figure 5.3: Relative error of the eigenvalue location via the hybrid method for over-soliton profile with 5 eigenvalues,  $q_{\text{over}}$ ,  $A = 5.25$ ,  $L = 20$ . The bubble size shows the inverse runtime (a smaller bubble means a longer run and vice versa), runtime changes from 3.27 s to 286 s.

Table 5.2: Normalised runtimes (in milliseconds per sample) of eigenvalues evaluation for multisoliton profile for iterative and contour integrals methods, implemented standalone, first row and column, and for hybrid algorithm ( $A = 5.25$ ,  $N_q = 2^{10}$ ).

	NR	secant	Steffensen	Muller	Sidi
	19.18	3.56	3.53	3.56	10.22
<b>DL</b>	7.68	7.3	76.83	7.31	35.33
7.8					
<b>aDL</b>	3.88	3.44	71.75	3.41	35.86
3.28					

the BO algorithm. The NR, Muller and secant algorithms perform fairly similar in terms of runtime, but the Muller method displays a worse accuracy. If we combine the NR and secant methods with any of the integral methods, then the utilization of both the DL and aDL approaches give similar results. It means that both the DL and aDL methods supply a sufficiently good initial approximation to reach our desired zero.

## 5.4 Accuracy and channel noise tolerance

### 5.4.1 Discretisation and additive noise influence

The main parameter that impacts the accuracy of a given NFT method, is the temporal domain discretisation step  $\Delta\tau$ . Having in mind the optical communications application, I tested the contour integrals approach using the sequence of profiles with randomly chosen eigenvalues from the given constellations that were suggested for NFT-based transmission [50]. Here we use two kinds of constellations depicted in the inset of fig. 5.5.

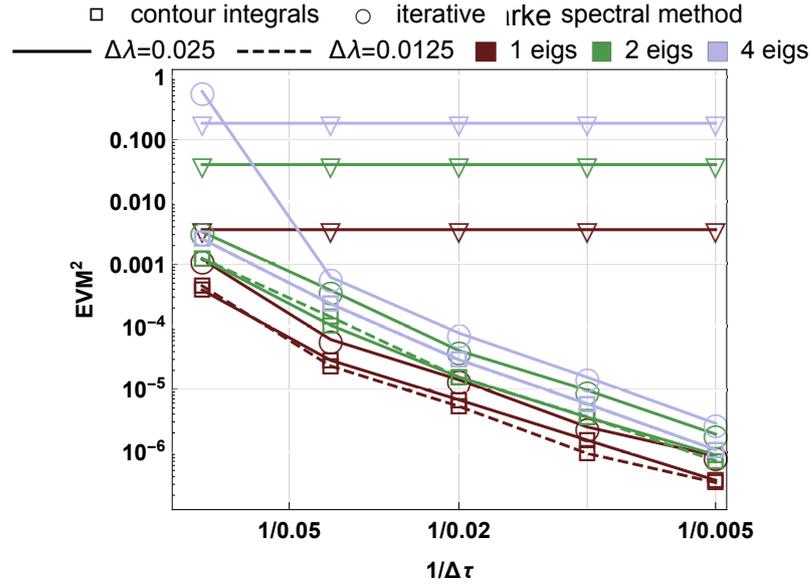


Figure 5.4:  $\text{EVM}^2$  as a function of time domain sampling  $\Delta\tau$  and contour sampling  $\Delta\lambda$ , for noiseless profiles with different number of eigenvalues. Results for the spectral method correspond to  $n = 51$  samples in the spectral domain.

I examine the performance of the contour integrals approach for single solitons, 2-eigenvalues and 4-eigenvalues profiles. I use the Darboux method, subsection 3.3.3, to generate such states with the arbitrary chosen norming constants and compute the respective eigenvalue portrait in the perfect back-to-back scenario (deterministic case, fig. 5.4) and in back-to-back adding a AWGN to the profile in the temporal domain, fig. 5.5. As a metric of computational accuracy I use EVM, see subsection 2.6.3 where  $\langle \cdot \rangle$  designates the averaging over the sequence of profiles and over the eigenvalues in the profiles. The normalisation over the mean value of constellation points appears after centring of the constellation diagram.

For solving the ZSP I use here BO method, known to have second-order accuracy and often superlative performance compared to other options, see section 4.2. The contour integrals  $s_p$  are computed using the trapezoidal rule along the contour. However, it is noteworthy that the error of the ZSP solution ( $a(\lambda)$  evaluation at each  $\lambda$ ) is  $O(\Delta\tau^2)$ , while the contour integration is carried out inside the spectral domain giving the spectral error of the order of  $O(\Delta\lambda^2)$ , such that choosing the consistent number of subintervals along the contour would require some further study. In this work, I manually adjusted the sampling of the contour and ZSP to balance error caused by both sources and to not oversample neither the integration contour  $C$ , not ZSP time domain.

The back-to-back tests are performed for different temporal domain discretisation step  $\Delta\tau$ , and the tests for noise-corrupted profiles were performed studying the accuracy in dependency on SNR, defined as a fraction of signal and noise powers, irrespectively of the solitonic nature of our signals. The SNR value was altered through AWGN with varying noise power in the time domain, because the eigenvalue (i. e. constellation structure) uniquely defined the signal power, so I do not have the freedom to manipulate it.

In figs. 5.4, 5.5 it is seen the reduction of EVM with the decrease of temporal domain discretisation step  $\Delta\tau$  and with the increase of SNR. By changing the contour sampling  $\Delta\lambda$ , one can manipulate the numerical accuracy of the contour integrals approach and get advantage comparably to the iterative method. The accuracy is typically better for the eigenvalues located along the imaginary axis (the constellation, corresponding to the filled circles in the inset of fig. 5.5). The number of eigenvalues slightly influences the value of

numerical error.

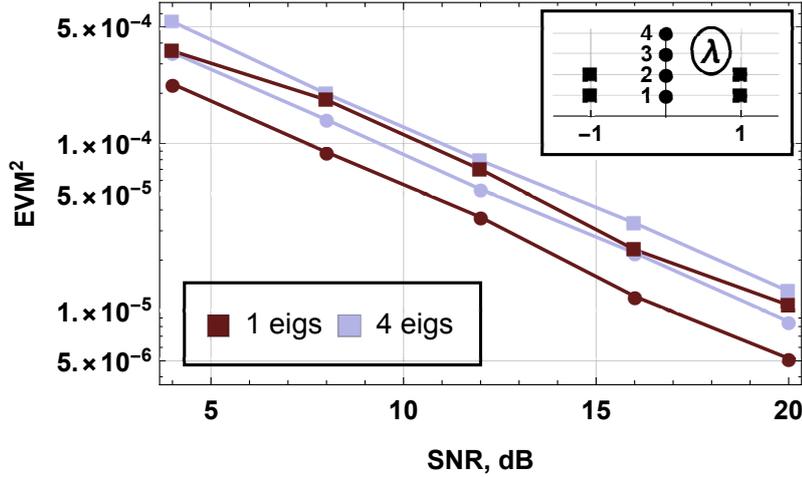


Figure 5.5:  $EVM^2$  for different eigenvalues constellations (circles - along the imaginary axis, squares - symmetrically around the imaginary axis) for a single eigenvalue and a 4-eigenvalues solution in dependence on AWGN SNR (time domain is sampled with  $\Delta\tau = 0.007$ , contour is sampled with  $\Delta\lambda = 0.0125$ ).

One of the advantages of contour integrals approach is its capability to easily evaluate a degenerate eigenvalue. To demonstrate this, I use an example of the degenerate breather [67]: a limiting case of two-eigenvalue soliton, when the eigenvalues approach each other. In the case of purely imaginary degenerate eigenvalue  $\lambda = ix$ , we can use the profile given in [67]:

$$q_{\text{deg}}(\tau) = \frac{8ix(2\tau x \sinh(2x\tau) - \cosh(2x\tau))}{8x^2\tau^2 + \cosh(4x\tau) + 1}. \quad (5.8)$$

In the ideal case, both algorithms should return two equal eigenvalues, and it would mean that  $a(\lambda)$  has a double zero at that point.

We observe that similarly to the non-degenerate case, the accuracy of grid search is worse than that for the iterative and contour integrals approaches (see fig. 5.6). Decreasing  $\Delta\lambda$ , we can improve the accuracy for both grid search and contour integrals algorithms. The error for contour integrals is generally smaller and degrades faster with the increase of  $1/\Delta\tau$  than the error for the iterative algorithm.

The actual behaviour of the contour integrals and iterative methods for the degenerate eigenvalues search is compared in fig. 5.7 (the results for all eigenvalues are given in the coordinate system with the origin at the eigenvalue, the red square). None of the used numerical methods locates both eigenvalues at exactly the same point even for a noiseless profile. The contour integrals approach gives the symmetrical points around the actual eigenvalues, and their deviations are smaller than those for the iterative method. The initial guess point also influences the accuracy of the iterative method's result, see the scattered points in the lower part of fig. 5.7).

The performance of the contour integrals approach using the degenerate breather with added noise, in dependence on the SNR was also studied, fig. 5.8. The results obtained are similar to those presented in fig. 5.5 earlier.

However, one of the important findings is that the accuracy of the computation of the eigenvalues increases with the magnitude of the degenerate eigenvalue. The latter is related to signal power. Therefore, I anticipate that the applicability of the contour integrals approach will be even more important for the high power signal processing. The paramount goal of NFT processing is to operate in the highly-nonlinear regime where conventional linear techniques fail.

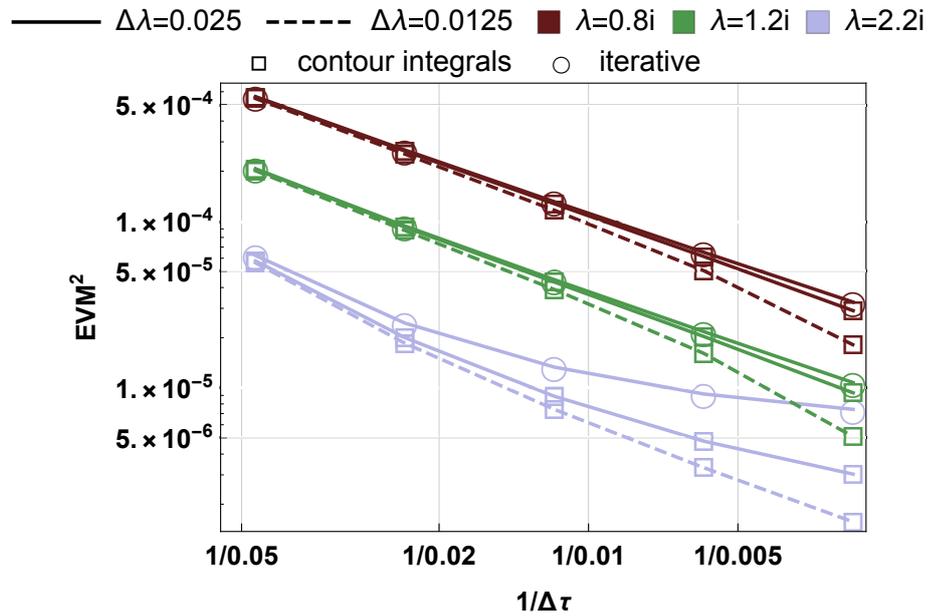


Figure 5.6:  $EVM^2$  behaviour for degenerate eigenvalue evaluation via different numerical method using noiseless degenerate breather profile, eq. (5.8), in dependence on the time domain discretisation  $\Delta\tau$ .

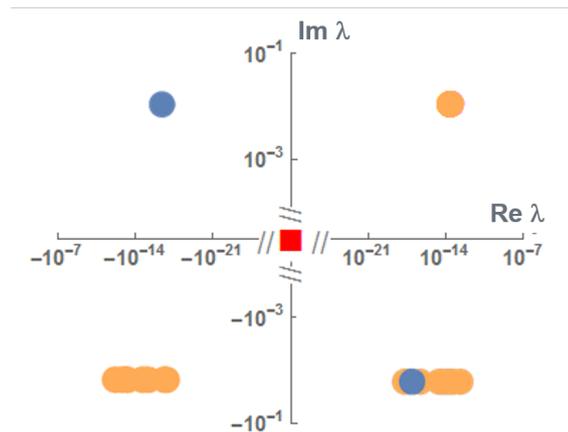


Figure 5.7: Deviation of numerically evaluated eigenvalues of the noiseless degenerate breather profile ( $\lambda = 1.2i$ ,  $\Delta\tau = 0.023$ ,  $\Delta\lambda = 0.0125$ ), computed via contour integrals and via iterative method.

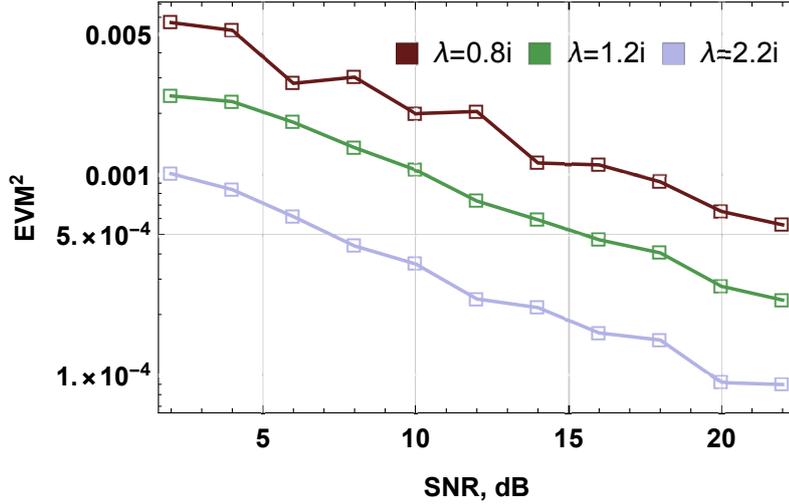


Figure 5.8:  $EVM^2$  of the degenerate breather with AWGN eigenvalue evaluation in dependence on AWGN SNR (time domain is sampled with  $\Delta\tau = 0.003$ , contour is sampled with  $\Delta\lambda = 0.0125$ ).

#### 5.4.2 Closely located eigenvalues

One of the advantages of the contour integrals approach is the ability to easily and simultaneously evaluate closely located eigenvalues. Evaluation of the multiple eigenvalues by means of iterative algorithms requires additional adjustment (e.g., one given in algorithm 1). Moreover, when the function  $a(\lambda)$  is redefined to remove already located eigenvalue, it causes 0/0 problems in numerics, which would reveal especially for close located eigenvalues. On the contrary, the contour integrals are free of these disadvantages. To prove this, I tested accuracy (expressed in terms of the relative error, see subsection 2.6.3) of both methods depending on the spacing between eigenvalues. In this test system, two eigenvalues are located symmetrically w.r.t. imaginary axis with gap distance  $\Delta$ , so their values are  $\lambda_{\pm} = \pm\Delta/2 + \Im\lambda$ . I tried various imaginary parts of the eigenvalues to reveal its influence as well.

The error of the iterative method rises dramatically when  $\Delta$  decreases, whilst the contour integrals behaviour is more stable. For large gap size (right part of the plot) another property of the DL method reveals: when the zero is located close to the integration contour, the numerical error of the method increases when it directly lies on the contour this method is inapplicable. In [30] the authors reported inverse proportionality of the error to the shortest distance from the contour to the zeros inside. That is the reason why we observe the increase of the relative error for larger  $\Delta$  for contour integrals approach.

#### 5.4.3 Influence of noise in the NLS channel

The fact that the contour integrals are uniquely determined by the set of eigenvalues corresponded to the waveform (cf. eq. (5.1)), makes them invariant during the signal propagation according to NLS. Also, the way of contour integrals computation allows more predictable runtime (because of the fixed contour) and allows avoiding possible singularities in the complex  $\lambda$ -plane. Both these facts motivate to consider the contour integrals as an alternative of the eigenvalue(s) modulation in NFT-based communication.

In this subsection, I perform analytical derivations of the contour integrals tolerance to practical deviations from the integrability, basing on the eigenvalue dynamics, and having in mind  $N_s$ -eigenvalue waveform. To the best of my knowledge, the only estimation of the eigenvalue statistics in the noisy channel is the theory of single soliton propagation in NLS

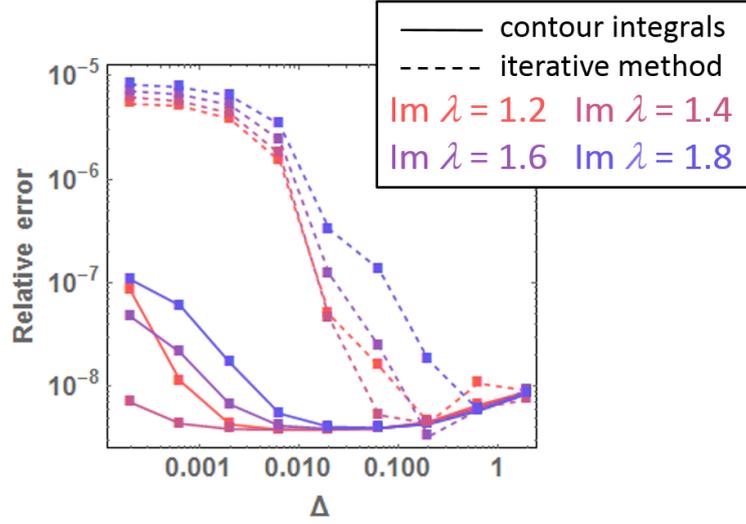


Figure 5.9: Relative error of numerically evaluated eigenvalues by contour integrals and iterative algorithms for closely located eigenvalues, symmetrically relative to the imaginary axis with the gap  $\Delta$ :  $\lambda_{\pm} = \pm\Delta/2 + \Im\lambda$ .

channel with AWGN, provided, e.g., in [55]. If one considers eq. (2.15) with additive noise term in r.h.p. (see subsection 2.6.4 for notations), then a single soliton with amplitude  $A_0$  after propagating over the distance  $\zeta$ , has eigenvalue's components with normal (Gaussian) distribution with the following standard deviations:

$$\sigma_A^2 = \Gamma_0 A_0 \zeta, \quad \sigma_{\Omega}^2 = \Gamma_0 A_0 \zeta / 3. \quad (5.9)$$

I use this estimation to derive the standard deviation of the contour integrals values, assuming that in the multieigenvalue case, the jitters of each solitonic component are independent and coincide with the single-eigenvalue case:

$$\delta\lambda_j \sim \frac{1}{2} (\delta\Omega_j + i \delta A_j), \quad |\delta\lambda_j| = \sqrt{\frac{\Gamma_0 A_j \zeta}{3}}. \quad (5.10)$$

Assuming also that noise is small perturbation above the signal and consider the first order of the perturbation theory for the contour integrals values deviation:

$$s_p = s_{0,p} + \delta s_p = \sum_{j=1}^{N_s} (\lambda_{0,j} + \delta\lambda_j)^p \approx s_{0,p} + p \sum_{j=1}^{N_s} \delta\lambda_j \lambda_{0,j}^{p-1} \quad (5.11)$$

where subscript 0 means an unperturbed value.

The goal of this derivation is to compare the relative deviation of the eigenvalue and the relative deviation of the contour integrals. The estimation of the single soliton jitters allow writing for the multieigenvalue profile:

$$\left| \frac{\delta\lambda_j}{\lambda_j} \right| = \frac{2\sqrt{\Gamma_0 \zeta}}{\sqrt{3}} \sqrt{\frac{A_j}{\Omega_j^2 + A_j^2}}, \quad (5.12)$$

whilst the deviation of the contour integrals is estimated as:

$$\begin{aligned}
 |\delta s_p| &= p \left| \sum_{j=1}^{N_s} \delta \lambda_j \lambda_{0,j}^{p-1} \right| \leq p \sum_{j=1}^{N_s} |\delta \lambda_j| |\lambda_{0,j}|^{p-1}, \\
 \left| \frac{\delta s_p}{s_p} \right| &\leq p \frac{(\max_j |\delta \lambda_j / \lambda_j|)^p}{(\min_j |\delta \lambda_j / \lambda_j|)^{p-1}}.
 \end{aligned} \tag{5.13}$$

Substituting the expression for the eigenvalue from soliton parameters, one can get:

$$\left| \frac{\delta s_p}{s_p} \right| \leq \frac{2p\sqrt{\Gamma_0\zeta} \sum_{j=1}^{N_s} \sqrt{A_j} \left( \Omega_j^2 + A_j^2 \right)^{\frac{p-1}{2}}}{\sqrt{3} \sum_{j=1}^{N_s} \left( \Omega_j^2 + A_j^2 \right)^{p/2}}. \tag{5.14}$$

One relative variance in the terms of another can be expressed as:

$$\left| \frac{\delta s_p}{s_p} \right| \leq p \frac{\sum_{j=1}^{N_s} A_j^{p/2} |\delta \lambda_j / \lambda_j|^{-(p-1)}}{\sum_{j=1}^{N_s} A_j^{p/2} |\delta \lambda_j / \lambda_j|^{-p}}. \tag{5.15}$$

In this dependency no extra entries of  $\zeta$ , so, the jitters dynamics with propagation of the signal along the fibre is expected to be the same.

Having in mind the communication application, I consider that all eigenvalues of the multieigenvalue profile are located within a circle in the upper half-plane with radius  $\Lambda_0$  and margins not less than  $x$ . In order to get an advantage of using the contour integrals for coding/decoding the information, it is reasonable to require that the relative deviation of the contour integrals values in the presence of noise are smaller than their eigenvalues' counterpart. For  $s_p$  one can get the following upper bound estimation:

$$\left| \frac{\delta s_p}{s_p} \right| \leq p \frac{2\sqrt{\Gamma_0\zeta}}{\sqrt{3}} \sqrt{2\Lambda_0} \sqrt{\frac{(x+2)^{p-1}}{x^p}}, \tag{5.16}$$

when for  $\lambda_j$ , it lies in the range:

$$\frac{2\sqrt{2\Lambda_0\Gamma_0\zeta}}{\sqrt{3}} \sqrt{\frac{1}{x+2}} \leq \left| \frac{\delta \lambda_j}{\lambda_j} \right| \leq \frac{2\sqrt{2\Lambda_0\Gamma_0\zeta}}{\sqrt{3}} \sqrt{\frac{1}{x}}. \tag{5.17}$$

The relative (normalised) deviation of the contour integrals increases with the order  $p$  faster than linearly, whilst for the eigenvalues, this estimation reveals no dependency on the number of eigenvalues, because we assume that solitonic modes interact independently and with the noise only. For the specific number of the eigenvalues, embedded to the profile, one can derive the condition when the normalised deviation of the contour integrals (the highest order one, when  $p = N_s$ ) would have advantage comparably to the eigenvalues deviation. In this case:

$$N_s \Lambda_0 \left( 1 + \frac{2}{x} \right)^{N_s-1} < 1, \quad \Lambda_0 < \frac{1}{N_s} \left( \frac{x}{x+2} \right)^{N_s-1}. \tag{5.18}$$

According to this formula, for higher  $N_s$  the allowed area of the eigenvalues in the complex plane squeezes more tightly. For  $x = 10\%$  and  $N_s = 8$ , it returns  $\Lambda_0 \approx 7 \cdot 10^{-11}$ , which is unrealistic for numerics.

To illustrate and test the provided estimations, I simulate the propagation of 2-4-eigenvalues pulses (with no continuous spectrum) through the noisy NLS channel on various propagation distances. The pulses were generated by means of Darboux Transform (DT) [10, 11]. Because the eigenvalues are invariant, it was sufficient to truncate the

solitonic tails at the Tx with large enough temporal window. The set of the eigenvalues was  $\lambda_1 = -1 + 0.8i$ ,  $\lambda_2 = 1i - 0.5$ ,  $\lambda_3 = 1.4i + 0.7$  and  $\lambda_4 = 1.8i + 1.1$ . The truncation points were chosen symmetrically and to have outside the dedicated window signal samples which do not exceed  $10^{-6}$  of the maximal signal amplitude. In solitonic scales, the simulated propagation distances lie within the range  $\zeta \in [0.5, 5]$  solitonic lengths. The noise intensity in the NLS channel corresponds to the normalised AWGN power spectral density  $\Gamma_0 = 2 \cdot 10^{-5}$ .

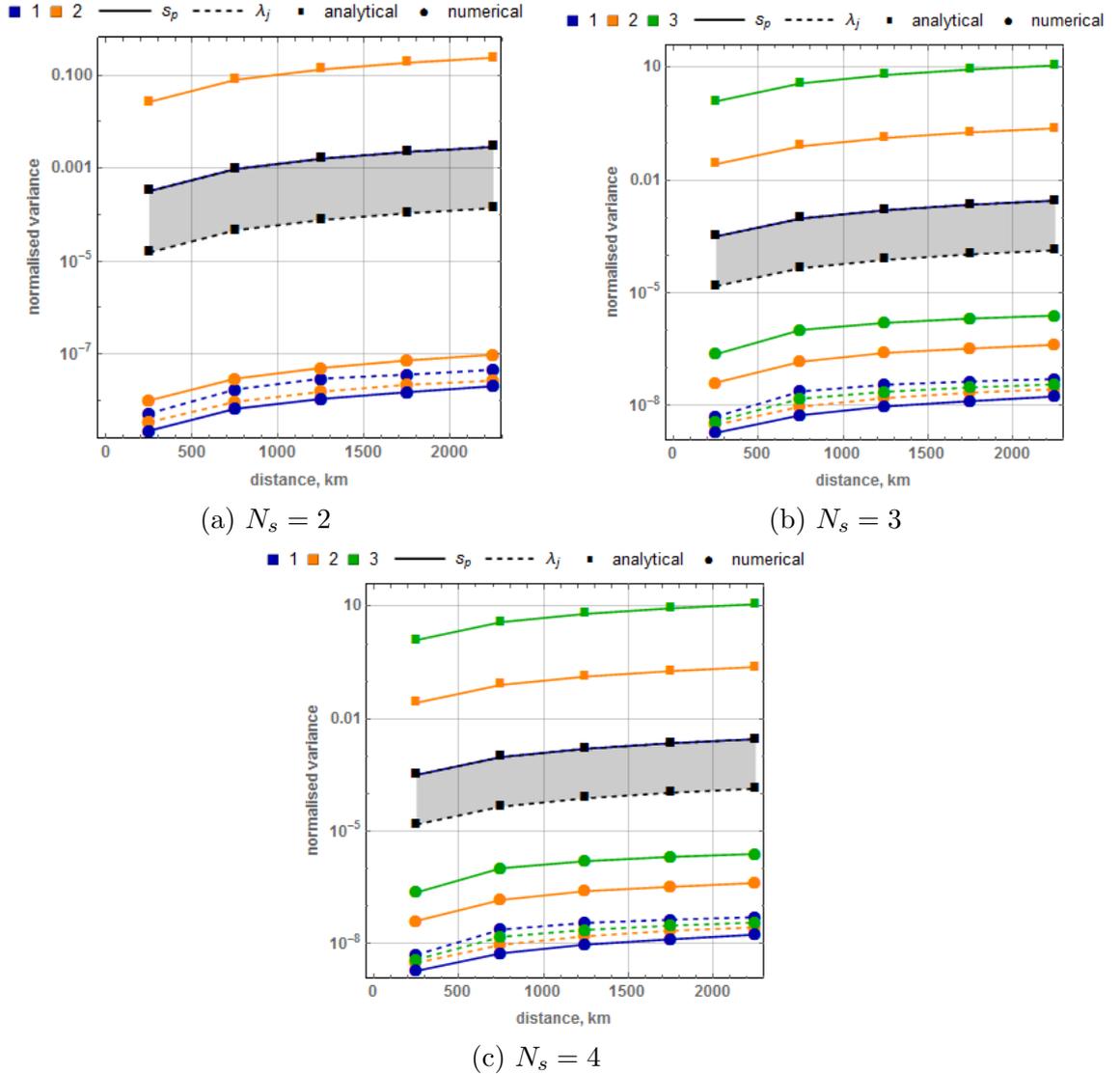


Figure 5.10: Normalised variance of the eigenvalues and contour integrals for multieigenvalue profiles after the propagation through NLS channel with AWGN associated with the amplifier noise.

It is seen from the plots in fig. 5.10 that the contour integrals deviating much more noticeably in the presence of noise. For higher order of  $s_p$  the (normalised) deviation increases. The analytical estimation of the eigenvalue deviation, given as black shadowed area according to the range from eq. (5.17), overestimated the actual deviation, obtained from numerical simulations. However, for more solitonic modes, the analytical estimation for the contour integrals deviations approached the numerical curves, which makes this estimation more applicable for multisoliton profiles.

## 5.5 Contour integrals approach discussion

In this chapter, several findings within the contour integrals approach for eigenvalues localisations are presented. This method was developed to replace the existent techniques for the discrete spectrum computation, which are known for suffering from the increasing numerical accuracy and unreliable performance. The contour integrals localise all eigenvalues inside a given ROI on one instance with manageable accuracy and runtime, scaling with the perimeter of ROI. One of the most important advantages of the contour integral is that it can successfully deal with cases when other methods fail, such as closely located or coinciding (degenerative) eigenvalues. The more practical case for the communication application, when several eigenmodes may be incorporated as data carriers, is the effective and reliable performance of the contour integrals in the multisoliton solutions.

On one hand, the contour integrals are combinable with all high-accuracy numerical methods for ZSP (such as BO, AL), on the other hand, it can be further refined with help of iterative methods, reaching noticeable accuracy for few iterations. This trick allows reducing efforts for the accuracy of initial contour integrals estimation.

The analytical study of noise influence on the contour integral values estimates their deviation to growing for higher-order integrals, However, because of the assumptions during the estimation, the perturbation of higher-order integrals is underestimated. Therefore, involving simulations to the analysis of noise tolerance of the contour integrals computation provides a more accurate picture of performance.

The approach can be easily generalised to the identification of the main spectrum point for the finite-band case. It also can be formulated as a problem of zero search in the complex plane, similarly combined with the solution of ZSP. In the systems, developed and studied during this research, the built-in methods for main spectrum points appear to be sufficient, and making a negligible contribution of the transmission quality. However, further developing of RHP-based communications towards higher data rates, so, introducing more bands in finite-band solution, may add a need to the applications of contour integrals for the periodic Nonlinear Fourier transform (PNFT) main spectrum.

## Chapter 6

# Optical communication schemes based on NFT

### 6.1 Overview

The concept of NFT-based communication requires that the channel model is effectively given by means of NLS. The realistic simulations usually consider fibre loss and amplification noise as well as significant phenomena, happen to the signal in the optical fibre. However, if one considers the pure integrable NLS as a leading order approximation and loss with ASE noise to be the next order correction, it is a way to use the nonlinear modes in NF domain as information carriers.

While the model is considered to be true, the information is invariant and not spoiled by the nonlinearity like in the conventional optical communications. In fig. 6.1, the reader can find the general concept of the NFT-based communication, literally inherits the procedure of the solution of the initial-value problem of NLS, fig. 3.1. The NF modes evolve independently, while the signal propagates in the fibre according to NLS. However, we note that deviations of the model from pure NLS, such as path averaging, loss, amplification, ASE noise, lead to the effective interaction of these NF modes. Study of the noise arising in NF domain took some part of my PhD study, leading to publications [J5, , CC1, CC4, CC6, CC7].

In the following sections, I will provide some details of the transmission systems. I use capital notations  $P$ ,  $B$ ,  $Z$ ,  $T$  for signal power, bandwidth, propagation distance, and, if applicable, period, measured in physical units. As the approach utilised some normalisation of NLS to the dimensionless form (see subsection 2.2.2), determined by the single normalisation scale  $T_0$ , I used subscript  $X_{\text{norm}}$  to designate corresponding normalised values, e.g.  $P_{\text{norm}}$  is normalised signal power.

### 6.2 Conventional NFT transmission

The variety of degrees of freedom in vanishing boundary conditions case allows us to employ different combinations of them for information transmission. The most widely acknowledged approach is to use continuous spectrum  $r(\lambda)$  to modulate information in the continuous waveform, and (optionally) incorporate discrete spectrum components. The NFT approach allows mathematically explicitly linearise the NLS channel, so the Kerr nonlinearity is considered here as a valuable part of the model, but not as an undesirable deviation from linearity. This inspired the research in the topic from the conceptual side, giving birth to eigenvalues communication [54], nonlinear inverses synthesis [84, 85, 117]. Moreover, this approach is generalisable to the dual-polarisation case as well. There have

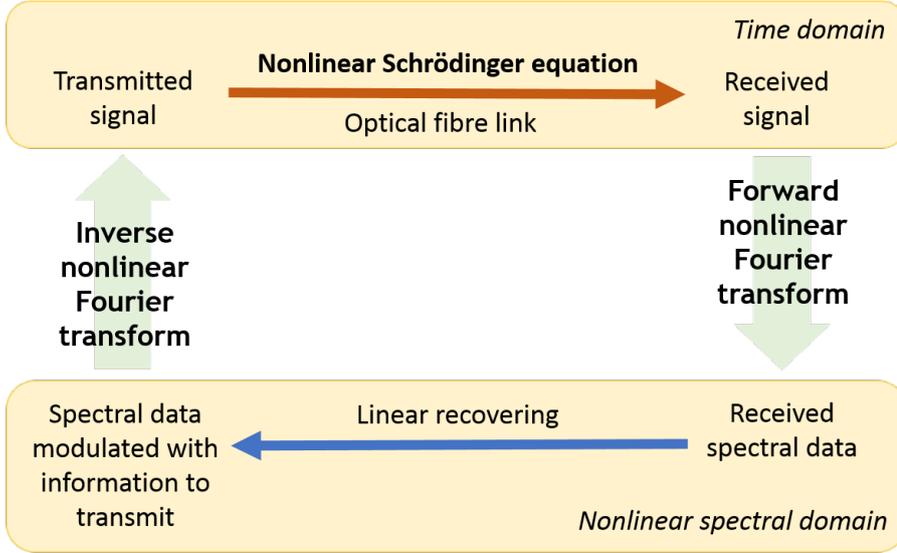


Figure 6.1: General concept of NFT-based communication.

already been a number of experimental confirmations that the NFT-based techniques do indeed have the potential for mitigation of nonlinear penalties [46, 40, 82, 83, 81].

In this thesis, I omit pure eigenvalues communication, as during my work I did not directly use it in conventional NFT-based transmission context. However, to some extent, it is discussed in the following subsections.

### 6.2.1 $r$ -modulation

In this subsection, I provide the simulations of the transmission system, based on modulation of continuous spectral data only. Meanwhile, in these simulations we test the eligibility of the fast NFT methods [140, 142] for the processing at the Tx and the Rx. The results presented here were submitted for ECOC 2018, but have been rejected.

We simulate the transmission through the SMF link for both EDFA (with path averaging) and Raman amplification, see subsection 2.1.4. The normalisation time parameter is  $T_0 = 50$  ps and, for the fairness of the comparison, the sampling step for both algorithms is  $\Delta\tau = 0.2$ , as it enters as a parameter into the numerical algorithms defining their performance. The information is first encoded in a orthogonal frequency-division multiplexing (OFDM) signal with 128 sinc carriers satisfying the Nyquist condition, using 16-QAM constellation. The schematic of our setup is shown in fig. 6.2.

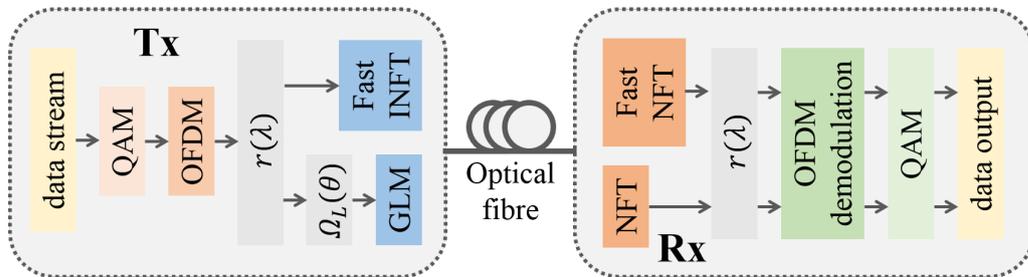


Figure 6.2: The principal scheme of  $r$ -modulation transmission system.

During noise-free transmission, the performance of two systems is affected only by the power deteriorating the numerical processing: the Q-factor is significantly larger for small powers, fig. 6.3. We explain this behaviour with numerical errors, including those due

to the truncation to finite support, which increases for larger powers. In the presence of noise, the high power degradation is still present, but small power signals are distorted by noise, fig. 6.4. Thus, there is an optimal power for transmission. Importantly, we observed that the quality of the transmission of the fast methods, where function  $r(\lambda)$  is naturally involved, is better than that of the conventional method with (de-)modulation of  $\Omega_L(\theta)$  from GLM, see eq. (3.57).

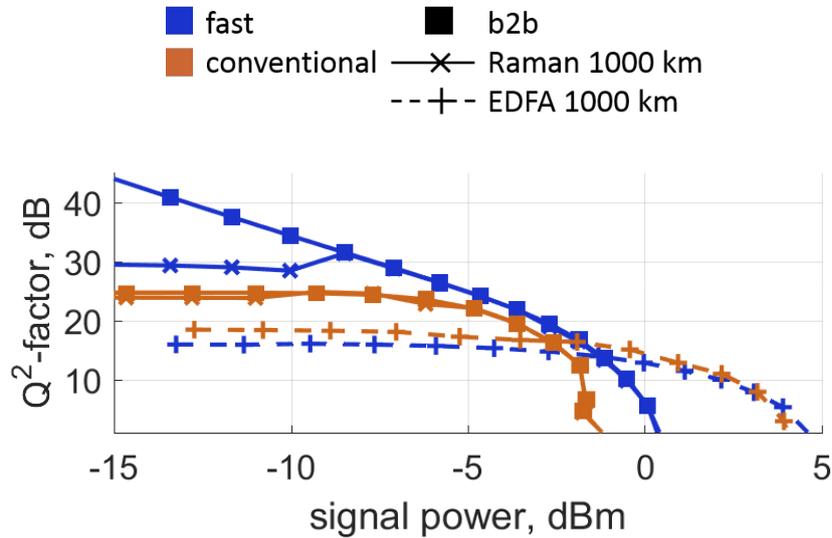


Figure 6.3:  $Q^2$  factor vs. power for  $r$ -modulation in the noiseless channel.

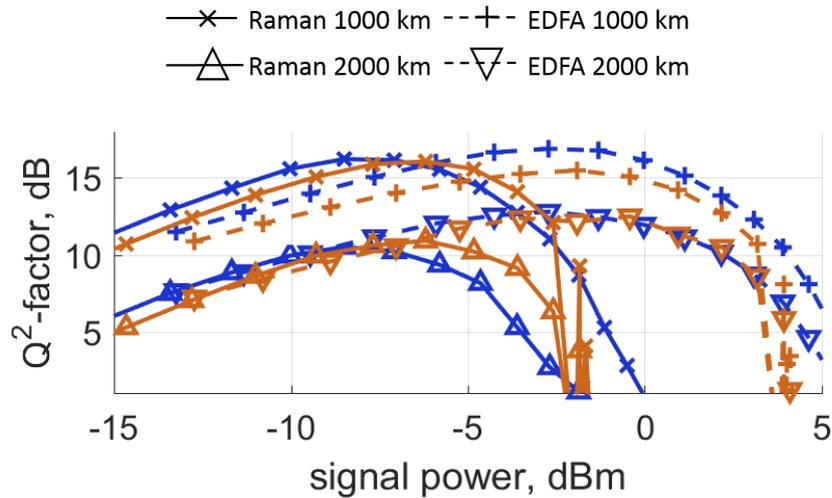


Figure 6.4:  $Q^2$  factor vs. power for  $r$ -modulation in the presence of ASE noise.

We also found that the optimal Q-factor slightly depends on the distance and consistently decreases for longer links, see fig. 6.5.

In terms of time consumption of each processing option, fig. 6.6, the profit of Rx processing on using fast methods (the fast AL against and conventional AL realisation) is more noticeable, than for the Tx, where the fast INFT [137] competes with the GLM solution method [38]. The gap between the conventional and fast INFT is expected to widen for a larger number of samples and more sophisticated implementation.

Generally speaking, the  $r$ -modulation, in particular, the so-called nonlinear inverse synthesis approach [83, 84, 85] is the most direct implementation of the NFT concept to communication purposes. There are several principal barriers, which do not allow effective

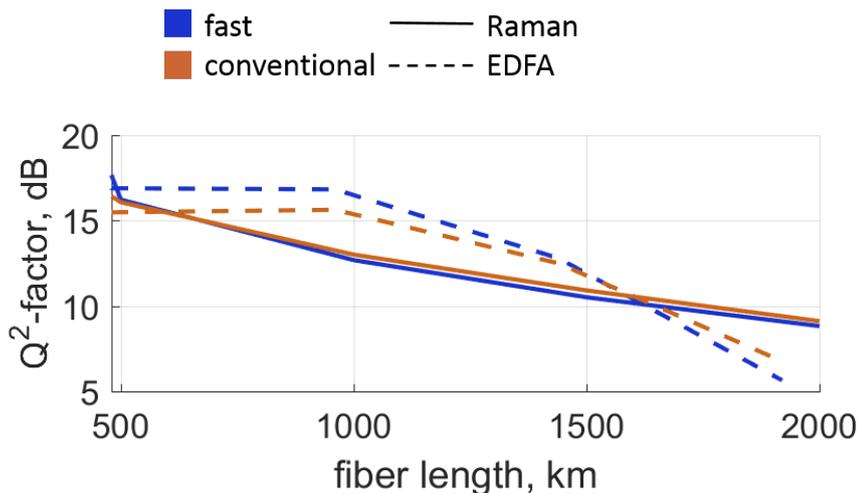


Figure 6.5: The optimal  $Q^2$  factor for different propagation distances in  $r$ -modulation.

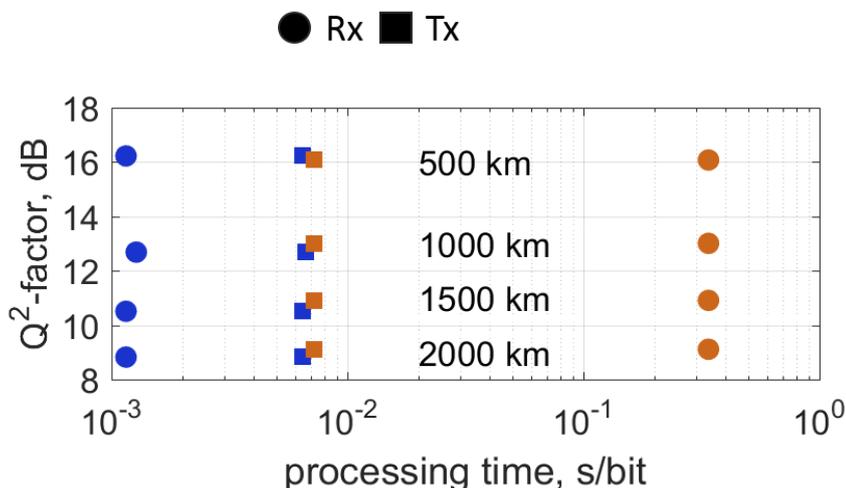


Figure 6.6: The processing time (per bit of information) required to reach the maximum  $Q^2$  at the Tx and Rx for different transmission distances for the systems based on fast and conventional  $r$ -modulation.

realisations. In fact, that is the nonlinear (and not always explicit) coupling of time-domain parameters with  $\lambda$  domain. The natural intensity scale in NF space is spectrum energy, eqs. (3.22) and (3.23), whereas for conventional telecommunication power, bandwidth and duration are the most relevant. If it is roughly possible to link energy and power, the bandwidth is not predictable. The time occupation is also related to energy, but there is no clear correspondence, just for the given modulation one may observe that higher energy produces the appearance of heavier signal tail [117]. Also, there is no analogy of (Nyquist–Shannon) sampling theorem for NF spectrum, so we cannot properly discretize the system without penalty in terms of accuracy, and we cannot manipulate over the required and the available number of samples in time and  $\lambda$  domain. All these properties and observations to the fact that any NFT-based transmission system requires preliminary tuning of parameters, depending on the desirable propagation distance, amplification, data rate, etc. Therefore the design of such systems is a quite bulky process. For example, the simulation of Raman and EDFA amplification for the system above, keeping all other output parameters, took noticeable efforts and time of this work. Also, the fact that even the most straightforward implementation of NFT idea causes a huge level of resistance

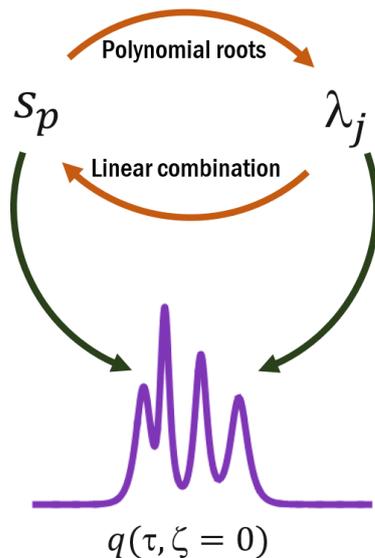


Figure 6.7: Schematics of one-to-one correspondence between NLS profile, eigenvalues and contour integrals.

and lack of understanding among the community, despite the clearly claimed advantage in nonlinearity mitigation, puts more sophisticated approaches to an even weaker position.

### 6.3 Contour integrals communication

This section contains some preliminary research in further developing of the contour integrals within IST, particularly, using their values to carry the information via NLS channel. The results in this section have not been presented before in any form.

The developing of the concept of employing the contour integrals values as information carriers should take into account the advantages of this approach in comparison with currently used iterative methods. One of the main is that contour integrals allow more predictable complexity and evaluation of all eigenvalues simultaneously. Since the set of the eigenvalues uniquely determines the multieigenvalue waveform, and simultaneously determines values of  $s_p$  through Newton's identities, eq. (5.3). It means that there is a one-to-one correspondence between the set of  $s_p$  and the waveform, and moreover, it conserves during the signal propagation in NLS channel, see fig. 6.7.

Principally, the idea of the modulation into the contour integrals values moves the additional processing stage (computation of the eigenvalues from the contour integrals) from the Rx stage on the Tx one. It makes this processing free of the influence of the noise gathered during the practical propagation in the optical fibre, balances the numerical complexities of the algorithms on the Tx and the Rx. The conceptual scheme of conventional multieigenvalue modulation and introduced contour integrals modulation are given in fig. 6.8.

This suggestion requires detailed comprehensive testing before being implemented in order to achieve the maximal advantage of this approach comparably to the well-developed multieigenvalue communication [52, 54].

If one starts with the values of the contour integrals  $s_p$ , taken randomly from some constellation, it is not obvious that the resulting eigenvalues  $\lambda_j$  found as polynomial zeros via eqs. (5.2)-(5.3), would lie in the upper half-plane. To ensure this, I suggest an additional transformation. Assuming that all  $s_p$  lie in the complex plane within the circle of radius  $S$ , the corresponding set of the eigenvalues  $\lambda_j$ , because of the polynomial relations, eq. (5.1),

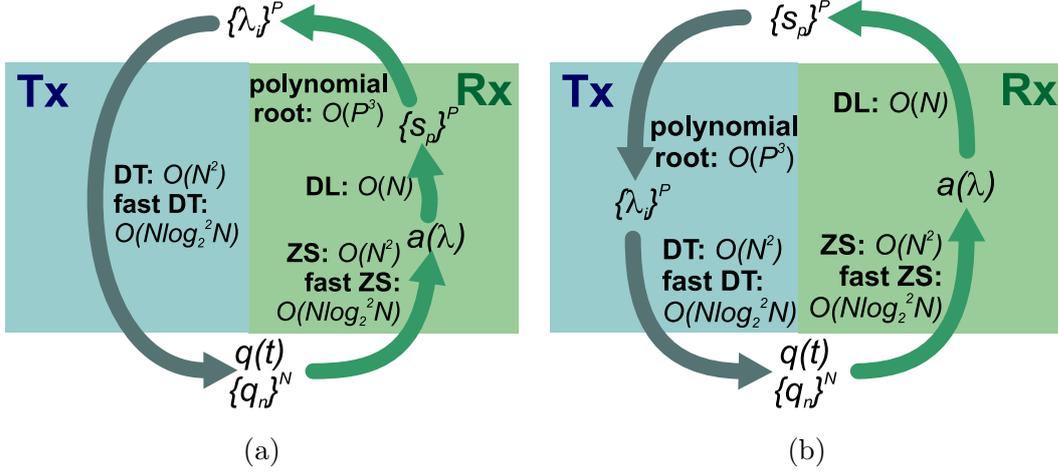


Figure 6.8: The schemes of (a) the multi-eigenvalue communication concept and (b) the contour integrals communication concept.

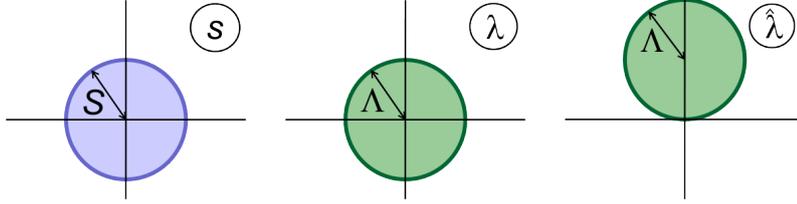


Figure 6.9: Regions of values of  $s_p$ ,  $\lambda_j$  and  $\hat{\lambda}_j$  in the corresponding complex planes for comparison of multi-eigenvalue and contour integrals communication. This adjustment ensures that the eigenvalues supplied to DT have positive imaginary parts.

would lie in the deterministic circle of radius  $\Lambda$ . One can shift the entire set up, i.e.  $\hat{\lambda}_j = i\Lambda + \lambda_j$ , to ensure that its interior now lies completely in the upper half-plane (see fig. 6.9).

The link between values of  $S$  and  $\Lambda$  should be derived, aiming the compact localisation of the eigenvalues in order to get a reasonable estimation of the resulting signal power and bandwidth, related to real and imaginary parts of the eigenvalues. It is unrealistic to assume that, especially for many eigenvalues, all  $s_p$  are limited by the same circle radius  $S$ , because in this case the limiting value would be large and the estimation of  $\Lambda$  is overvalued. So, I assume that each order  $s_p$  lies in the circle with radius  $S_p$ , increasing exponentially as  $|s_p| \leq S_p = N_s C^p$ , for  $N_s$ -eigenvalues profile and for some scale parameter  $C$ . Then we need the intermediate estimation for the limiting value of Newton's identities  $|\kappa_p| \leq K_p$ , found using the Cramer's rule and the Hadamard's inequality, as  $\kappa_p$  are given as solutions of the linear equations system, eq. (5.3):

$$\kappa_p = \frac{\det \mathbf{S}_p}{\det \mathbf{S}} = \frac{\det \mathbf{S}_p}{N_s!} \leq \frac{\prod_l \|\mathbf{S}_l\|}{N_s!}, \quad (6.1)$$

Note that here I employ the fact that the equations system matrix  $\mathbf{S}$  is triangular, and the matrix  $\mathbf{S}_p$  is the same matrix with the substituted  $p$ -th column by the r.h.s. of the system, which is  $(s_1, \dots, s_{N_s})^T$ . The simplification leads to the following expression:

$$K_p = \frac{N_s^{N_s+1} C}{N_s!} \frac{\sqrt{C^{2N_s} - 1} \sqrt{C^{N_s^2+3N_s+2} (C^{-2(N_s+1)}; C^2)_{N_s+1}}}{(C^2 - 1)^{N_s/2} (C^{2N_s+2} - 1) \sqrt{p^2 (C^2 - 1) + N_s^2 C^2 (C^{2N_s-2p} - 1)}}, \quad (6.2)$$

where the notation  $(a; q)_n$  means  $q$ -Pochhammer symbol [44]. Applying this for the estimation of the polynomial roots from their coefficients, given by Affane-Aji, Agarwal and

Govil in [3], one can get:

$$\Lambda = \frac{1}{2} \left( 1 + CN_s + \sqrt{(1 - CN_s)^2 + 4 \frac{CN_s \sqrt{C^{2N_s} - 1}}{\sqrt{C^2 - 1}} \left( 1 + \frac{N_s^2 C^2}{C^2 - 1} (C^{2N_s - 2} - 1)^{-1/2} \right)} \right), \quad (6.3)$$

assuming that the scaling factor  $C \leq 1/\sqrt{2}$ .

The specific values of these estimations allow adjust the scaling parameter  $C$  to the desired signal power and bandwidth, defined by  $\Lambda$  and NLS normalisation (see subsection 2.2.2).

## 6.4 $b$ -modulation NFT transmission

*Note that in this section I use  $k$  as a spectral NF parameter notation (instead of  $\lambda$ ).*

Apart from the straightforward modulation of the function  $r(k)$ , which can be treated as a generalisation of (conventional) Fourier spectrum modulation, there is a more advanced and novel approach, introduced by Wahls in [138]. Its idea lies in the fact that both auxiliary scattering functions  $a(k)$  and  $b(k)$ , eq. (3.17) satisfy the following integral representations [37, 138]:

$$\begin{aligned} a(k) &= 1 + \int_0^\infty \alpha(\theta) e^{ik\theta} d\theta, \\ b(k) &= \int_{-\infty}^\infty \beta(\theta) e^{ik\theta} d\theta \end{aligned} \quad (6.4)$$

In case of the finite time-support of the potential  $q(\tau)$ , say,  $q(|\tau| > L/2) = 0$ , the infinite limit in integrals can be replaced with  $2L$  interval, so eq. (6.4) can be interpreted as that the  $b(k)$  waveform is explicitly determined by its Fourier image  $\beta(\theta)$ , defined over a finite domain  $\theta$  or, in other words, function  $b(k)$  is band-limited. The opportunity to generate from band-limited  $b(k)$  finitely supported  $q(\tau)$  allows to get a more efficient time-domain occupation compared to ‘‘conventional’’ NFT-based systems employing the continuous NF spectrum modulation [84, 135], the property which translates into a better system performance, see the explicit comparison in [47]. We, however, note that the  $b$ -modulation concept is not different in its set-up from other NFT-based methods. Thus, within this approach, each ‘‘supersymbol’’ generated from the  $b$ -modulated profile must be appended with zero-padding guard intervals in the time domain, which are equal in duration to the dispersion-induced memory. However, the duration of the ‘‘supersymbol’’ itself is shorter in comparison to the conventional continuous NF spectrum modulation where, typically, the time-domain waveforms develop ‘‘a tail’’ [117].

The utilisation of this property of NFT-generated profiles, operating with band-limited  $b(k)$  profiles, has been aimed at resolving one of the principal challenges in the NFT-based communication: to attain explicit control over the temporal duration of generated signals at the Tx side [47, 86, 145]. The latter property allows us to pack our data better inside a given time-bandwidth volume and, thus, to reach potentially higher SE values. In particular, the highest data rate (400 Gbps for approximately 1000 km of the transmission distance) reported so far for the NFT-based transmission method was achieved with a modified variant of  $b$ -modulation (in the dual-polarisation case) [146] and has recently been confirmed experimentally [147].

My research in this topic and results achieved are presented as the conference contributions [C1, C4], where the performance of  $b$ -modulated transmission system is paired with eigenvalues communication; and as journal publication [J4], where the formal proof of the backward correspondence between band-limited  $b(k)$  and finitely supported  $q(\tau)$  is provided, together with characterisation of the allowances in discrete spectrum. In works

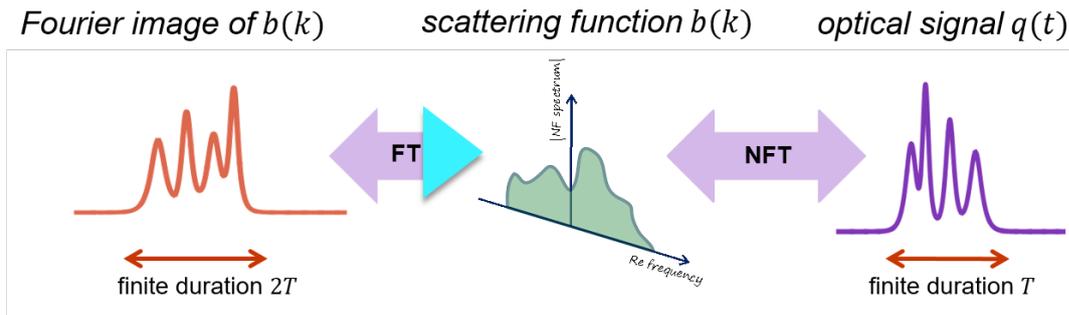


Figure 6.10: Illustration of one-to-one correspondence between bandlimited function  $b(k)$  and the corresponding finitely supported time-domain profile  $q(\tau)$ , highlighting the contribution of the current study (necessity condition in violet arrows, when proven here sufficiency condition in blue arrows).

[C1, C4] we composed the discrete spectrum components atop the  $b$ -modulated profile, but to keep the width of the solitary component in the time domain sufficiently thin. Such a composition ensures that the considerable portion of the overall resulting signal does not spread beyond the initial extent of the  $b$ -modulation-generated profile. However, this approach did not provide truly localised signals. In work [J4] we suggested the procedure of the identification of the specific eigenvalues, which maintain the exact localisation of the signal.

#### 6.4.1 Formal justification of signal and auxiliary $b$ -function correspondence

For the completeness of exposition here, we note that in the original work by Wahls [138], where the  $b$ -modulation concept was introduced, the problem of a complete characterization of  $b(k)$  in the case when time-limited signals  $q(\tau)$  support bound states (i.e. containing a non-zero discrete NF spectrum part), was formulated as an open question. In the follow-up study [47], a necessary condition for the possibility to have bound states keeping the same  $b(k)$  was stated and the analogy with the linear operator of the Lax pair representation for KdV equation was mentioned. Portinari [115] pointed out that  $b(k)$  being band-limited is both necessary and sufficient for a finite support in the KdV case if there are no bound states. For the NLS case (i.e., for the Zakharov-Shabat system as a spectral problem), the uniqueness of the determination of a time-limited  $q(\tau)$  in (3.5) from  $b(k)$  is discussed in Ref. [125] in the absence of bound states. The arguments of Portinari were recently carried over the NLS case in [136], where the characterisation problem has also been addressed in the presence of eigenvalues, by exploiting the spectral properties of one-sided signals.

The sufficiency of the condition for receiving the exactly localised signal was not properly addressed before our publication [J4]. There we have applied the flexibility of RHP approach to  $b$ -modulated signals [72, 151], [J1] for solving the inverse scattering problem for ZSP (3.5). As an illustration of the performed contribution, refer to fig. 6.10.

By this way, we *characterize* time-limited signals having the same scattering coefficient  $b(k)$  showing, in particular, that it is possible to include the discrete nonlinear spectral components (solitons) into the  $b$ -modulation without violating the complete localisation of the respective time-domain profile. In the end, it is noteworthy that whereas the inclusion of additional solitary components can increase the overall power of the signal and, thus, improve the signal-to-noise ratio defined in the “traditional sense”, this, of course, does not necessarily mean that the transmission performance of the  $b$ -modulated system would

get better. The coupling of solitary modes to the members of the encoded continuous spectrum due to the deviation of the channel from the integrable NLS (e.g. due to the presence of noise, in the simplest scenario) should definitely affect and potentially degrade the performance of the system in hand [11], [CC6] though this question has not been studied in detail so far. From the other side, the possibility to embed solitons in the  $b$ -modulation renders additional flexibility and new opportunities for the design of such systems with improved characteristics.

Here we note some specific properties of IST of finitely supported  $q(\tau)$ , which appear in addition to the listed in section 3.1. Notice that for finite-extent  $q(t)$ ,  $\Psi(\tau, k)$  are the entire analytic functions of  $k \in \mathbb{C}$ . For the consistency of presentation, we give here a simple proof of the property from eq. (6.4) using the integral representations for the Jost solutions (cf. [125]).

The following proofs are organised as follows. First, we prove the necessary conditions, following the RHP formalism used later. This is a known property of scattering function and its proof can be found elsewhere [37, 138]. Then, we develop a sequence of supporting statements (lemmas, remarks, etc), and prove a sufficiency condition by means of RHP approach in theorem 6.4.6.

**Theorem 6.4.1.** *Let  $q \in L^1(\mathbb{R})$  be such that  $q(\tau) = 0$  for  $|\tau| > \frac{L}{2}$  for some  $L > 0$ . Then*

- $\Psi^{lr}(\tau, k) = e^{-ik\tau\sigma_3}$  for  $\tau < -\frac{L}{2}$ , and

$$\Psi^{lr}(\tau, k) = e^{-ik\tau\sigma_3} + \int_{-L-\tau}^{\tau} \mathbf{K}_1(\tau, \tau') e^{-ik\tau'\sigma_3} d\tau', \quad \tau > -\frac{L}{2}; \quad (6.5)$$

- $\Psi^{rl}(\tau, k) = e^{-ik\tau\sigma_3}$  for  $\tau > \frac{L}{2}$ , and

$$\Psi^{rl}(\tau, k) = e^{-ik\tau\sigma_3} + \int_{\tau}^{L-\tau} \mathbf{K}_2(\tau, \tau') e^{-ik\tau'\sigma_3} d\tau', \quad \tau < \frac{L}{2}. \quad (6.6)$$

Here  $\mathbf{K}_j(\tau, \cdot) \in L^1$ ,  $j = 1, 2$ , are some  $2 \times 2$  matrix functions.

*Proof of Theorem 6.4.1.* For any  $q(\tau) \in L^1(\mathbb{R})$ , the integral representation for  $\Psi$  has the form [154]:

$$\Psi(\tau, k) = e^{-ik\tau\sigma_3} + \int_{-\infty}^{\tau} \mathbf{K}_1(\tau, \tau') e^{-ik\tau'\sigma_3} d\tau'. \quad (6.7)$$

Assuming for a moment that  $q(\tau)$  is smooth, i.e  $q(\tau) \in C^1(\mathbb{R})$ , substituting (6.7) into (3.5), and applying  $\int_{-\infty}^{\infty} (\cdot) e^{ik\tau''} dk$ , it follows that  $\mathbf{K}_1(\tau, \tau')$  satisfies the system of equations:

$$\begin{aligned} \mathbf{K}_1(\tau, \tau) - \sigma_3 \mathbf{K}_1(\tau, \tau) \sigma_3 &= \mathbf{Q}(\tau), \\ \frac{\partial \mathbf{K}_1(\tau, \tau')}{\partial \tau} + \sigma_3 \frac{\partial \mathbf{K}_1(\tau, \tau')}{\partial \tau'} \sigma_3 - \mathbf{Q}(\tau) \mathbf{K}_1(\tau, \tau') &= 0, \quad \tau' < \tau, \end{aligned} \quad (6.8)$$

where  $\mathbf{Q}(\tau)$  is given as:

$$\mathbf{Q}(\tau) = \begin{pmatrix} 0 & q(\tau) \\ -q^*(\tau) & 0 \end{pmatrix}. \quad (6.9)$$

Decomposing  $\mathbf{K}_1$  into the diagonal and off-diagonal parts,  $\mathbf{K}_1^d$  and  $\mathbf{K}_1^o$ , respectively,

$$\mathbf{K}_1 = \mathbf{K}_1^o + \mathbf{K}_1^d, \quad (6.10)$$

Eq. (6.8) then reduces to

$$\begin{aligned} \mathbf{K}_1^o(\tau, \tau) &= \frac{1}{2} \mathbf{Q}(\tau), \\ \frac{\partial \mathbf{K}_1^o(\tau, \tau')}{\partial \tau} - \frac{\partial \mathbf{K}_1^o(\tau, \tau')}{\partial \tau'} - \mathbf{Q}(\tau) \mathbf{K}_1^d(\tau, \tau') &= 0, \quad \tau' < \tau, \\ \frac{\partial \mathbf{K}_1^d(\tau, \tau')}{\partial \tau} + \frac{\partial \mathbf{K}_1^d(\tau, \tau')}{\partial \tau'} - \mathbf{Q}(\tau) \mathbf{K}_1^o(\tau, \tau') &= 0, \quad \tau' < \tau. \end{aligned} \quad (6.11)$$

Now changing the variables as  $\xi = \tau + \tau'$ ,  $\eta = \tau - \tau'$ , and  $\tilde{\mathbf{K}}(\xi, \eta) := \mathbf{K}_1(\tau, \tau')$ , with

$$\frac{\partial \tilde{\mathbf{K}}}{\partial \xi} = \frac{1}{2} \left( \frac{\partial \mathbf{K}_1}{\partial \tau} + \frac{\partial \mathbf{K}_1}{\partial \tau'} \right), \quad \frac{\partial \tilde{\mathbf{K}}}{\partial \eta} = \frac{1}{2} \left( \frac{\partial \mathbf{K}_1}{\partial \tau} - \frac{\partial \mathbf{K}_1}{\partial \tau'} \right), \quad (6.12)$$

system (6.11) reduces to the following one:

$$\begin{aligned} \tilde{\mathbf{K}}^o(\xi, 0) &= \frac{1}{2} \mathbf{Q} \left( \frac{\xi}{2} \right), \\ \frac{\partial \tilde{\mathbf{K}}^o(\xi, \eta)}{\partial \eta} &= \frac{1}{2} \mathbf{Q} \left( \frac{\xi + \eta}{2} \right) \tilde{\mathbf{K}}^d(\xi, \eta), \quad \eta > 0, \\ \frac{\partial \tilde{\mathbf{K}}^d(\xi, \eta)}{\partial \xi} &= \frac{1}{2} \mathbf{Q} \left( \frac{\xi + \eta}{2} \right) \tilde{\mathbf{K}}^o(\xi, \eta), \quad \eta > 0. \end{aligned} \quad (6.13)$$

In turn, eq. (6.13) reduces to an integral equation of Volterra type. Indeed, integrating (6.13) we have:

$$\begin{aligned} \tilde{\mathbf{K}}^o(\xi, \eta) &= \tilde{\mathbf{K}}^o(\xi, 0) + \frac{1}{2} \int_0^\eta \mathbf{Q} \left( \frac{\xi + \eta'}{2} \right) \tilde{\mathbf{K}}^d(\xi, \eta') d\eta' \\ &= \frac{1}{2} \mathbf{Q} \left( \frac{\xi}{2} \right) + \frac{1}{2} \int_0^\eta \mathbf{Q} \left( \frac{\xi + \eta'}{2} \right) \tilde{\mathbf{K}}^d(\xi, \eta') d\eta', \\ \tilde{\mathbf{K}}^d(\xi, \eta) &= \frac{1}{2} \int_{-\infty}^\xi \mathbf{Q} \left( \frac{\xi' + \eta}{2} \right) \tilde{\mathbf{K}}^o(\xi', \eta) d\xi'. \end{aligned} \quad (6.14)$$

Substituting the second expression from eq. (6.14) into the first one, we arrive at a single integral equation:

$$\begin{aligned} \tilde{\mathbf{K}}^o(\xi, \eta) &= \frac{1}{2} \mathbf{Q} \left( \frac{\xi}{2} \right) + \frac{1}{4} \int_0^\eta \mathbf{Q} \left( \frac{\xi + \eta'}{2} \right) \times \\ &\times \int_{-\infty}^\xi \mathbf{Q} \left( \frac{\xi' + \eta'}{2} \right) \tilde{\mathbf{K}}^o(\xi', \eta') d\xi' d\eta'. \end{aligned} \quad (6.15)$$

Now notice that for  $\xi < -L$ , we have  $\mathbf{Q} \left( \frac{\xi}{2} \right) = 0$ , and, thus, eq. (6.15) becomes a homogeneous Volterra integral equation (in the domain  $\xi < -L$ ,  $\eta > 0$ ), the unique solution of which is 0. Therefore,  $\tilde{\mathbf{K}}(\xi, \eta) \equiv 0$  for  $\xi < -L$ ,  $\eta > 0$ , and, thus,  $\mathbf{K}_1(\tau, \tau') = 0$  for  $\tau + \tau' < -L$ . The general case of  $q \in L^1$  follows further by approximating  $\mathbf{Q}$  in (6.15) by smooth functions.

Similarly,  $\Psi^{rl}(\tau, k)$  has the representation

$$\Psi^{rl}(\tau, k) = e^{-ik\tau\sigma_3} + \int_\tau^\infty \mathbf{K}_2(\tau, \tau') e^{-ik\tau'\sigma_3} d\tau', \quad (6.16)$$

where, actually,  $\mathbf{K}_2(\tau, \tau') = 0$  for  $\tau + \tau' > L$ , which can be proven by following similar arguments as above and taking into account that  $\mathbf{Q}(\xi/2) = 0$  for  $\xi > L$ .  $\square$

Our main result consists in the characterization of signals  $q(\tau)$  having the  $b$ -coefficients in the form of the Fourier transform of a function with limited (bounded) support, i.e. the inverse of the property below.

**Corollary 6.4.2.** *1. In this case, the associated scattering functions  $a(k)$  and  $b(k)$  have the following integral representations:*

$$a(k) = 1 + \int_0^{2L} \alpha(\theta) e^{ik\theta} d\theta, \quad b(k) = \int_{-L}^L \beta(\theta) e^{ik\theta} d\theta, \quad (6.17)$$

with some  $\alpha(\theta) \in L^1(0, 2L)$ ,  $\beta(\theta) \in L^1(-L, L)$ .

2. For  $\tau > L/2$ ,  $b(k)e^{2ik\tau} \rightarrow 0$  as  $k \rightarrow \infty$  for  $k \in \mathbb{C}_+$ , and  $b^*(k^*)e^{-2ik\tau} \rightarrow 0$  as  $k \rightarrow \infty$  for  $k \in \mathbb{C}_-$ .
3. For  $\tau < -L/2$ ,  $b^*(k)e^{-2ik\tau} \rightarrow 0$  as  $k \rightarrow \infty$  for  $k \in \mathbb{C}_+$ , and  $b(k)e^{2ik\tau} \rightarrow 0$  as  $k \rightarrow \infty$  for  $k \in \mathbb{C}_-$ .

Indeed, setting  $\tau = -L/2$  in the scattering matrix relation (3.14), and using eq. (6.6) and the fact that  $\Psi^{lr}(-\frac{L}{2}, k) = e^{ik\frac{L}{2}\sigma_3}$ , it follows that  $a$  and  $b$  have the representations in form of eq. (6.17), where  $\alpha(\theta) = (\mathbf{K}_2)_{(22)}(-\frac{L}{2}, \theta - \frac{L}{2})$  and  $\beta(\theta) = -(\mathbf{K}_2)_{(21)}(-\frac{L}{2}, \frac{L}{2} - \theta)$ . Items 2 and 3 directly follow from eq. (6.17).

The set of spectral data determining uniquely  $q(t)$ , is (conventionally) characterised assuming that  $a(k) \neq 0$  for  $k \in \mathbb{R}$  and all zeros of  $a(k)$  in  $\mathbb{C}_+$  are simple; consequently, the number of these zeros is finite, and  $|b(k)| < 1$  for all  $k \in \mathbb{R}$ .

For the following derivation we use the flexibility of RHP approach for the vanishing case, see corresponding preliminaries in subsection 3.3.4.

With these assumptions, the characteristic spectral data consist of  $b(k)$ ,  $k \in \mathbb{R}$  and the discrete set  $\{k_j, c_j\}_1^{N_s}$  (given, in general, independently of  $b(k)$ ; particularly, it can be empty), where  $k_j$  with  $\Im k_j > 0$ ,  $j = 1, \dots, N_s$ , are the zeros of  $a(k)$ , and  $\{c_j\}_1^{N_s}$  are the associated norming constants. Moreover, the inverse mapping can be described as follows [37]:

1. Given  $b(k)$  and  $\{k_j\}_1^{N_s}$ , construct  $a(k)$  in accordance with (3.25) for  $k \in \mathbb{R}$ :

$$a(k) = \prod_{j=1}^{N_s} \frac{k - k_j}{k - k_j^*} \exp \left\{ \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{\log(1 - |b(s)|^2)}{s - k} ds \right\}; \quad (6.18)$$

2. Define the reflection coefficient

$$r(k) := b(k)/a(k), \quad k \in \mathbb{R}; \quad (6.19)$$

3. Solve the corresponding RHP for the jump condition of  $k \in \mathbb{R}$ :

$$\mathbf{M}_+(\tau, k) = \mathbf{M}_-(\tau, k)\mathbf{J}(\tau, k), \quad k \in \mathbb{R}, \quad \mathbf{M}(\tau, k \rightarrow \infty) \rightarrow \mathbf{I}. \quad (6.20)$$

where

$$\mathbf{J}(\tau, k) = \begin{pmatrix} 1 + |r(k)|^2 & r^*(k)e^{-2ik\tau} \\ r(k)e^{2ik\tau} & 1 \end{pmatrix}. \quad (6.21)$$

The singularities of  $\mathbf{M}$  are characterised as follows:  $\mathbf{M}_{(1)}(\tau, k)$  has simple poles at  $\{k_j\}_1^{N_s}$  and  $\mathbf{M}_{(2)}(\tau, k)$  has simple poles at  $\{k_j^*\}_1^{N_s}$  such that the following residue conditions hold:

$$\text{Res } \mathbf{M}_{(1)}(\tau, k) \Big|_{k=k_j} = c_j e^{2ik_j\tau} \mathbf{M}^{(2)}(\tau, k_j), \quad (6.22)$$

$$\text{Res } \mathbf{M}_{(2)}(\tau, k) \Big|_{k=k_j^*} = -c_j^* e^{-2ik_j^*\tau} \mathbf{M}^{(1)}(\tau, k_j^*). \quad (6.23)$$

4. Having the RHP solved,  $q(\tau)$  can be obtained by

$$q(\tau) = 2i \overset{[i]}{\mathbf{M}}_{(12)}(\tau), \quad (6.24)$$

where the matrix  $\overset{[i]}{\mathbf{M}}(t)$  emerges from the large- $k$  development of  $\mathbf{M}(\tau, k)$ :

$$\mathbf{M}(\tau, k) = \mathbf{I} + \frac{\overset{[i]}{\mathbf{M}}(\tau)}{k} + O(k^{-2}), \quad k \rightarrow \infty. \quad (6.25)$$

**Remark 6.4.3.** In the framework of the direct scattering problem (given  $q(\tau)$ , determine the scattering data),  $\mathbf{M}(\tau, k)$  is related to Jost solutions of the ZSP, eq. (3.5) as follows

$$\mathbf{M}(\tau, k) = \begin{cases} \left( \frac{\Psi_{(1)}^{lr}(\tau, k)}{a(k)}, \Psi_{(2)}^{rl}(\tau, k) \right) e^{ik\tau\sigma_3}, & k \in \mathbb{C}_+, \\ \left( \Psi_{(1)}^{rl}(\tau, k), \frac{\Psi_{(2)}^{lr}(\tau, k)}{a^*(k)} \right) e^{ik\tau\sigma_3}, & k \in \mathbb{C}_-. \end{cases} \quad (6.26)$$

Let a function  $b(k) \neq 0$ ,  $k \in \mathbb{C}$ , be given such that

- (i)  $b(k) = \int_{-L}^L \beta(\theta) e^{ik\theta} d\theta$  with some  $\beta(\theta) \in L^1(-L, L)$ ;
- (ii) The function  $G(k) := 1 - b^*(k^*)b(k)$  has no zeros for  $k \in \mathbb{R}$  (or, equivalently,  $G(k) > 0$  for  $k \in \mathbb{R}$ ).

**Definition 6.4.4.** Define  $\mathcal{F}_b$  as the set of all  $q \in L^1$  such that the spectral functions  $b(k) = b(k; q)$  and  $a(k) = a(k; q)$  associated to  $q$  by the direct mapping through (3.5) and (3.14), satisfy the following conditions:

- (a)  $b(k; q)$  coincides with the prescribed function  $b(k)$  satisfying conditions (i) and (ii) above (particularly, we assume that  $a(k; q)$  is not zero for real  $k$ );
- (b) All zeros of  $a(k; q)$  for  $k \in \mathbb{C}_+$  are simple.

**Notation 6.4.5.** Denote by  $\mathcal{A}_b$  the set of all zeros of  $G(k)$ .

**Theorem 6.4.6.** 1. The set  $\mathcal{F}_b$  consists of infinitely many elements;

2. For any  $q \in \mathcal{F}_b$ ,  $q(\tau) = 0$  for  $|\tau| > L/2$ ;

3. Each particular  $q \in \mathcal{F}_b$  is uniquely specified by a finite subset  $\{k_j\}_1^{N_s}$  (including the empty set) of  $\mathcal{A}_b \cap \mathbb{C}_+$ . This subset constitutes the set of simple zeros of the spectral function  $a(k)$  in  $\mathbb{C}_+$  associated to this  $q$ , which can be expressed for  $k \in \overline{\mathbb{C}_+}$  by (6.18) in terms of  $b(k)$ ,  $k \in \mathbb{R}$  and  $\{k_j\}_1^{N_s}$ .

**Remark 6.4.7.** In the case  $b(k) \equiv 0$ , it follows from (3.25) that the associated  $a(k)$  has no zeros and thus, by (6.18),  $a(k) \equiv 1$ ; consequently,  $q(t)$  with such spectral data is the trivial one:  $q(t) \equiv 0$ .

The proof of the theorem above is based on using the flexibility of the RHP formalism: the same  $q(\tau)$  can be retrieved from the solutions of different RHPs; and one can proceed from one RHP to another (that produces the same  $q(\tau)$ ) by appropriately “transforming” the original RHP, e.g., factorizing the jump matrix and absorbing the factors into the solution of a new RHP problem, in specific domains of the complex plane.

First, notice that in our case  $b(k)$  is analytic in  $\mathbb{C}$  and thus the norming constants are determined by  $b(k)$  and  $a(k)$ :

$$c_j = \frac{b(k_j)}{\dot{a}(k_j)}, \quad (6.27)$$

where the overdot ( $\dot{\phantom{a}}$ ) means the derivative with respect to  $k$ .

*Proof that  $\mathcal{F}_b$  consists of infinitely many elements.* We notice that  $b(k)$  and  $G(k)$  are the entire functions of order at most 1 [87]. This follows from the following estimates:

$$|b(k)| = \left| \int_{-L}^L \beta(\theta) e^{ik\theta} d\theta \right| \leq e^{L|k|} \int_{-L}^L |\beta(\theta)| d\theta, \quad (6.28)$$

and

$$\begin{aligned} |G(k)| &\leq 1 + \left| \int_{-L}^L \int_{-L}^L e^{ik(\xi+\eta)} \beta(\xi) \overline{\beta(-\eta)} d\xi d\eta \right| \leq \\ &\leq 1 + e^{2L|k|} \int_{-2L}^{2L} |B(\theta)| d\theta, \end{aligned} \quad (6.29)$$

with  $B(\theta) = \int_{-L}^L \beta(\eta) \overline{\beta(\eta-\theta)} d\eta$  and  $B(\theta) \in L^1(-2L, 2L)$ . Actually,  $b(k)$  is the entire function of order 1 and type  $L$  provided  $\beta(\theta)$  does not vanish almost everywhere in any neighbourhood of  $L$  or  $-L$ , see [16, ss. 6.9.1]; similarly for  $G(k)$ .

Next, for an entire function of order at most 1, the Hadamard factorisation theorem, see [87, p. 26], implies that  $G(k)$  can be written as a product:

$$G(k) = k^m e^{p_1 k + p_2} \prod_{n=1}^M \left( 1 - \frac{k}{k_n} \right) e^{\frac{k}{k_n}}, \quad (6.30a)$$

or

$$G(k) = k^m e^{p_1 k + p_2} \prod_{n=1}^M \left( 1 - \frac{k}{k_n} \right), \quad (6.30b)$$

where  $m \geq 0$ ,  $p_j \in \mathbb{C}$ , and  $\{k_n\}_1^M$  with  $M \leq \infty$  are the zeros of  $G(k)$ .

Assume that  $M < \infty$ . Then it follows from (6.30) that

$$G(k) = e^{\tilde{p}_1 k + p_2} P(k), \quad (6.31)$$

with some  $\tilde{p}_1$  and  $p_2$ , where  $P(k)$  is a polynomial of degree  $m + M$ . On the other hand, we notice that for  $k = x \in \mathbb{R}$ ,  $b(x) = \int_{-L}^L \beta(\theta) e^{ix\theta} d\theta \rightarrow 0$  as  $x \rightarrow \pm\infty$  and thus  $G(x) \rightarrow 1$ . Combining this with  $G(x) = e^{\tilde{p}_1 x + p_2} P(x)$  evaluated as  $x \rightarrow +\infty$ , and as  $x \rightarrow -\infty$ , we conclude that  $\tilde{p}_1$  must be 0 and thus  $G(x) \equiv 1$ , which is in contradiction with  $b(k) \not\equiv 0$ .  $\square$

It follows from (3.25) that the set  $\mathcal{A}_b$  (determined by  $b(k)$  and symmetric w.r.t. the real axis) is a union of zeros of  $a(k; q)$  and  $a^*(k^*; q)$ . Consequently, all zeros of  $a(k; q)$  in  $\mathbb{C}_+$  associated with  $q \in \mathcal{F}_b$  (the eigenvalues of (3.5)) must be contained in  $\mathcal{A}_b \cap \mathbb{C}_+$ .

Let us choose any finite (particularly, it can be empty) subset  $\{k_j\}_1^{N_s}$  from  $\mathcal{A}_b \cap \mathbb{C}_+$ , construct  $a(k)$  for  $k \in \mathbb{C}_+$  by (6.18) (accordingly,  $a^*(k)$  is determined for  $k \in \mathbb{C}_-$ ), and determine  $r(k)$  for  $k \in \mathbb{C}_+$  by  $r(k) = b(k)/a(k)$  (cf. (6.19)) as well as  $r^*(k)$  for  $k \in \mathbb{C}_-$  by  $r^*(k) = b^*(k)/a^*(k)$ . Our main point is that using  $b(k)$ ,  $a(k)$ , and  $\{k_j, c_j\}_1^{N_s}$ , specified above, as the spectral data and the input to RHP, eqs. (6.20)–(6.23), one always arrives at such  $q(\tau)$  that  $q(\tau) = 0$  for  $|\tau| > L/2$ .

*Proof that  $q(\tau) = 0$  for  $\tau > L/2$ .* The proof is based on the transformation of the RHP (6.20)–(6.23):  $\mathbf{M} \mapsto \hat{\mathbf{M}}$ , suggested by the following algebraic factorization of  $\mathbf{J}$  in (6.21):

$$\mathbf{J}(\tau, k) = \begin{pmatrix} 1 & r^*(k)e^{-2ik\tau} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ r(k)e^{2ik\tau} & 1 \end{pmatrix}, \quad k \in \mathbb{R}. \quad (6.32)$$

Recall that the RHP (6.20)–(6.23) (particularly, the residue conditions (6.22), (6.23)) has been formulated under assumption that  $a(k)$  has no real zeros; otherwise (6.22), (6.23) are not correct.

For all  $\tau > L/2$ , define  $\hat{\mathbf{M}}(\tau, k)$  by

$$\hat{\mathbf{M}}(\tau, k) := \begin{cases} \mathbf{M}(\tau, k) \begin{pmatrix} 1 & 0 \\ -r(k)e^{2ik\tau} & 1 \end{pmatrix}, & k \in \mathbb{C}_+, \\ \mathbf{M}(\tau, k) \begin{pmatrix} 1 & r^*(k)e^{-2ik\tau} \\ 0 & 1 \end{pmatrix}, & k \in \mathbb{C}_-. \end{cases} \quad (6.33)$$

Notice that the triangular matrix factors in (6.33) considered (for any fixed  $t > L/2$ ) as functions of  $k$  in the corresponding half-planes are such that they are meromorphic in the respective half-planes and they approach  $\mathbf{I}$  exponentially fast as  $k \rightarrow \infty$ . The latter follows from Corollary 6.4.2, item 2, and the fact that  $a(k) \rightarrow 1$  as  $k \rightarrow \infty$  for  $k \in \mathbb{C}_+$ .

Particularly, this implies that the large- $k$  expansion for  $\hat{\mathbf{M}}(\tau, k)$  coincides with that for  $\mathbf{M}(\tau, k)$ , see (6.25), and thus, determining  $\hat{q}(\tau)$  from  $\hat{\mathbf{M}}(\tau, k)$  in the same way as  $q(\tau)$  is determined from  $\mathbf{M}(\tau, k)$ , see (6.24), we have:

$$\hat{q}(\tau) = q(\tau), \quad \tau > L/2. \quad (6.34)$$

On the other hand,  $\hat{\mathbf{M}}(\tau, k)$  can be characterized as the solution of the RHP with the trivial jump conditions: find  $\hat{\mathbf{M}}(\tau, k)$  analytic in  $\mathbb{C} \setminus \mathbb{R}$  and satisfying the properties

$$\begin{aligned} \hat{\mathbf{M}}_+(\tau, k) &= \hat{\mathbf{M}}_-(\tau, k), & k \in \mathbb{R}, \\ \hat{\mathbf{M}}(\tau, k) &\rightarrow \mathbf{I}, & k \rightarrow \infty. \end{aligned} \quad (6.35)$$

Indeed, in view of the factorization (6.32), the jump conditions across  $\mathbb{R}$  for  $\hat{\mathbf{M}}$  turn out to be trivial, and the statement is obvious in the case when  $a(k)$  has no zeros in  $\mathbb{C}_+$ . If  $a(k_j) = 0$  for some  $k_j \in \mathbb{C}_+$ , we evaluate  $\hat{\mathbf{M}}_{(1)}(\tau, k)$  as  $k \rightarrow k_j$  by using

$$\hat{\mathbf{M}}_{(1)}(\tau, k) = \mathbf{M}_{(1)}(\tau, k) - \frac{b(k)}{a(k)} e^{2ik\tau} \mathbf{M}_{(2)}(\tau, k) \quad (6.36)$$

that follows from eq. (6.33). Taking into account (6.22), it then follows that, as  $k \rightarrow k_j$ ,

$$\begin{aligned} \hat{\mathbf{M}}_{(1)}(\tau, k) &= \frac{1}{k - k_j} \frac{b(k_j)}{\dot{a}(k_j)} e^{2ik_j\tau} \mathbf{M}_{(2)}(\tau, k_j) + O(1) \\ &- \left( \frac{b(k_j)}{\dot{a}(k_j)(k - k_j)} e^{2ik_j\tau} \mathbf{M}_{(2)}(\tau, k_j) + O(1) \right) = O(1). \end{aligned} \quad (6.37)$$

Therefore,  $\hat{\mathbf{M}}_{(1)}(\tau, k)$  has no singularity at  $k = k_j$ . Similarly for  $\hat{\mathbf{M}}_{(2)}(\tau, k)$  at  $k = k_j^*$ .

The trivial jump conditions across  $\mathbb{R}$  in eq. (6.35) imply that  $\hat{\mathbf{M}}(\tau, k)$  is, in fact, analytic in the whole complex plane. Also it is bounded at infinity; moreover, it approaches the identity matrix as  $k \rightarrow \infty$ . Then, by the Liouville theorem,  $\hat{\mathbf{M}}(\tau, k) \equiv \mathbf{I}$  and thus, in view of (6.24) and (6.25),  $\hat{q}(\tau) \equiv 0$ . Finally, in view of eq. (6.34), we have  $q(\tau) = 0$  for  $\tau > L/2$ .  $\square$

*Proof that  $q(\tau) = 0$  for  $\tau < -L/2$ .* Like above, the proof is based on the deformations of the (original) RHP, eqs. (6.20)-(6.23). But now it is convenient to do the appropriate transformation in two steps:  $\mathbf{M} \mapsto \tilde{\mathbf{M}} \mapsto \tilde{\tilde{\mathbf{M}}}$ .

Step 1:  $\mathbf{M} \mapsto \tilde{\mathbf{M}}$ . Define

$$\tilde{\mathbf{M}}(\tau, k) := \begin{cases} \mathbf{M}(\tau, k) \begin{pmatrix} a(k) & 0 \\ 0 & \frac{1}{a(k)} \end{pmatrix}, & k \in \mathbb{C}_+, \\ \mathbf{M}(\tau, k) \begin{pmatrix} \frac{1}{a^*(k^*)} & 0 \\ 0 & a^*(k^*) \end{pmatrix}, & k \in \mathbb{C}_-. \end{cases} \quad (6.38)$$

Then, it follows from eqs. (6.20) and (6.21) that  $\tilde{\mathbf{M}}(\tau, k)$  satisfies the following jump conditions for  $k \in \mathbb{R}$ :  $\tilde{\mathbf{M}}_+(\tau, k) = \tilde{\mathbf{M}}_-(\tau, k) \tilde{\mathbf{J}}(\tau, k)$ , where

$$\begin{aligned} \tilde{\mathbf{J}}(\tau, k) &= \begin{pmatrix} a^*(k) & 0 \\ 0 & \frac{1}{a^*(k^*)} \end{pmatrix} \begin{pmatrix} 1 + |r(k)|^2 & r^*(k^*) e^{-2ik\tau} \\ r(k) e^{2ik\tau} & 1 \end{pmatrix} \begin{pmatrix} a(k) & 0 \\ 0 & \frac{1}{a(k)} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ \frac{b(k)}{a^*(k^*)} e^{2ik\tau} & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{b^*(k^*)}{a(k)} e^{-2ik\tau} \\ 0 & 1 \end{pmatrix}. \end{aligned} \quad (6.39)$$

Step 2:  $\tilde{\mathbf{M}} \mapsto \check{\mathbf{M}}$ . The triangular factorization in (6.39) suggests introducing the second RHP deformation step, defining  $\check{\mathbf{M}}$  by

$$\check{\mathbf{M}}(\tau, k) := \begin{cases} \tilde{\mathbf{M}}(\tau, k) \begin{pmatrix} 1 & -\frac{b^*(k^*)}{a(k)} e^{-2ik\tau} \\ 0 & 1 \end{pmatrix}, & k \in \mathbb{C}_+, \\ \tilde{\mathbf{M}}(\tau, k) \begin{pmatrix} 1 & 0 \\ \frac{b(k)}{a^*(k^*)} e^{2ik\tau} & 1 \end{pmatrix}, & k \in \mathbb{C}_-. \end{cases} \quad (6.40)$$

Now notice that the triangular factors in (6.40) are again meromorphic in the respective half-planes and, in view of Corollary 6.4.2, item 3, they approach  $\mathbf{I}$  exponentially fast as  $k \rightarrow \infty$ . Consequently, for  $\check{q}(t)$  obtained from the large- $k$  asymptote of  $\check{\mathbf{M}}(\tau, k)$  by (6.24) and (6.25) we have:

$$\check{q}(\tau) = q(\tau), \quad \tau < -L/2. \quad (6.41)$$

On the other hand, reasoning as in the case  $\tau > L/2$ , we can show that  $\check{\mathbf{M}}(\tau, k)$  has no singularities in  $\mathbb{C} \setminus \mathbb{R}$  and thus  $\check{\mathbf{M}}(\tau, k)$  can be characterised as the solution of the (piecewise analytic) RHP (6.35) with trivial jump conditions. As above, this implies that  $\check{\mathbf{M}}(\tau, k) \equiv \mathbf{I}$  and thus  $\check{q}(\tau) \equiv 0$ . In view of eq. (6.41),  $q(\tau) = 0$  for  $\tau < -L/2$ , which completes the proof of theorem 6.4.6.  $\square$

Here we note that theorem 6.4.6 completes the necessity and sufficiency correspondence between the band-limited  $b(k)$  profile and the finitely supported  $q(\tau)$  signal. However, this approach introduces the upper limitation of the spectrum energy and therefore, bound the maximal achievable signal power. This motivates further research in broadening the available power range via utilising discrete NF spectrum. Note that alternative ways of overcoming the energy limitation are given in [47] via the optimised waveform (see more details in subsection 6.4.2), or in record-breaking works [86, 146, 147].

### 6.4.2 Localised b-modulated profiles containing solitons

According to the discussion above, in order to embed the discrete spectrum components into the signal generated via the  $b$ -modulation method, which would not destroy the localisation of the signal in the time domain, the embedded discrete eigenvalues,  $k_{\text{eig}} \in \{k_j\}_1^{N_s}$ , must satisfy the condition, following from eq. (3.25) and item 3 of theorem 6.4.6:

$$b^*(k_{\text{eig}})b(k_{\text{eig}}) = 1. \quad (6.42)$$

For the known expression for  $b(k)$ , as it occurs in the optical transmission tasks employing the  $b$ -modulation, we can numerically seek for such points in the upper complex half-plane of parameter  $k$ . Then, these points give us the location where we can place our solitary modes without destroying the complete localisation of the time domain signal.

In the  $b$ -modulation approach, the signal power is manipulated by scaling of the modulated waveforms. However, this adjustment is typically performed numerically because of the non-trivial dependency between  $b(k)$  and  $q(\tau)$  scalings. In particular, let  $b(k) = Au(k)$ , where  $u(k)$  is the waveform modulated in a known way independently of the desired signal power, and assume that we do not have a discrete spectrum. The signal energy, given by the expression through the nonlinear spectrum functions as in eq. (3.22), together with the time support value  $L$ , define the signal power (in normalised units):

$$P_q = \frac{\varepsilon_{\text{cont}}}{L} = -\frac{1}{L\pi} \int_{-\infty}^{\infty} \log(1 - |b(k)|^2) dk. \quad (6.43)$$

Thus, having defined the particular signal power and modulation type, we can calculate the scaling factor  $A$  and, therefore, further define the locus of ‘‘allowed’’ eigenvalues, i.e.

Table 6.1: The values of  $a_m$  for the flat-top waveform, eq. (6.45) (cited from [47]).

$a_0=1.00781249999087$	$a_1=2.01562499996723$	$a_2=2.01562499848123$
$a_3=2.01562428510123$	$a_4=2.01557690160615$	$a_5=2.01459671013285$
$a_6=2.00542418293614$	$a_7=1.95813292084616$	$a_8=1.80756640511884$
$a_9=1.49055821347783$	$a_{10}=1.03117157326193$	$a_{11}=0.563957100582878$
$a_{12}=0.229897459751809$	$a_{13}=0.064961507923051$	$a_{14}=0.0112874144984265$
$a_{15}=0.000905697614069561$		

the eigenvalues that would not destroy the exact localisation, exploiting the theoretical results from subsection 6.4.1.

Here I present the procedure for the generation of a  $b$ -modulated signal with embedded discrete NF eigenmodes, which conserves the signal localisation, employing two simple carrier waveforms that are commonly used within the  $b$ -modulation approach [47, 138], the Nyquist shape, i.e. the sinc function in the  $k$ -domain and rectangular profile in the corresponding Fourier-conjugated domain:

$$u_{\text{sinc}}(k) = \text{sinc}(kL) := \frac{\sin(kL)}{kL}, \quad (6.44)$$

and the flat-top window carrier function, introduced in [47] as a way of overcoming the  $b$ -modulated signal power constraint:

$$u_{\text{flat}}(k) = \sum_{m=0}^{15} a_m [\text{sinc}(kL - m) + \text{sinc}(kL + m)], \quad (6.45)$$

where  $L$  is the expected time-domain duration, and the coefficients  $a_m$  are listed in table 6.1. The studied waveforms used for the  $b(k)$  modulation with their corresponding Fourier images are given in fig. 6.11.

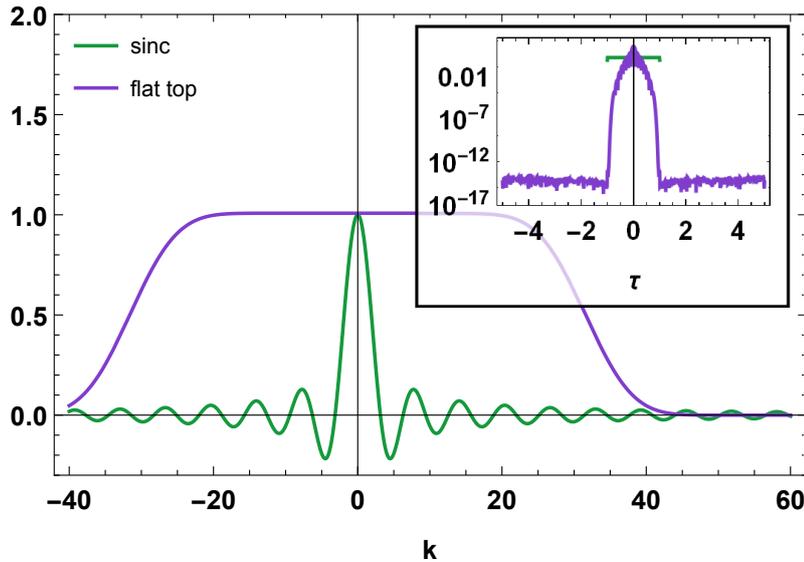


Figure 6.11: The waveforms used as the illustration of  $b$ -modulation method with their Fourier transforms.

Depending on the value of the scaling factor  $A$ , these functions have points in the complex plane of  $k$ , which can be used to implant our eigenvalues at, while keeping the

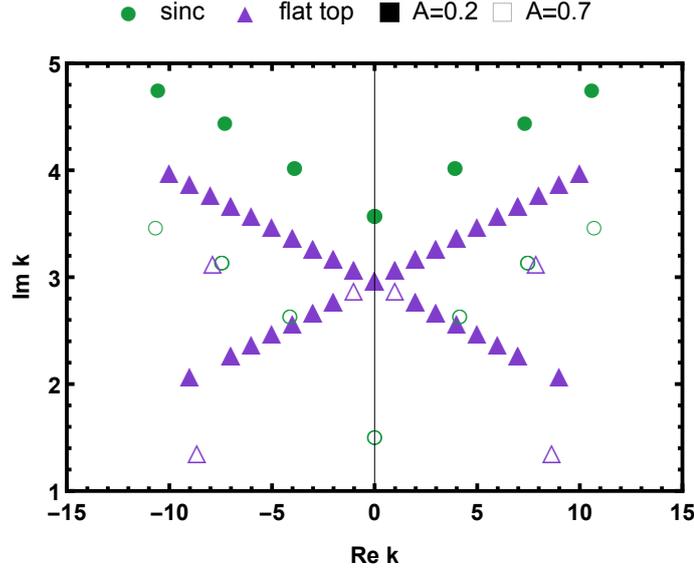


Figure 6.12: The points from  $\mathcal{A}_b$  available for placing the eigenvalues at, when keeping the exact localisation of  $b$ -modulated signal. Circles and triangles distinguish different  $b$ -shapes, while their filling identifies the level of the scaling factor  $A$ .

exact localisation of the resulting  $q$  profile. Of course, the numerical search cannot guarantee that we have identified all appropriate points, but at the moment we just need to find some of them to illustrate the idea. Typically, for communication purposes, we do not use eigenvalues with large real and/or imaginary parts. This occurs in view of the numerical issues associated with INFT computation for high-amplitude solitons, or since, e.g., a nonlinear eigenmode with a large  $\Re k_{\text{eig}}$  would rapidly escape from the dedicated time-window during the signal propagation. The numerically found set of points, which can be used as an eigenvalue locus for our studied  $b(k)$  waveforms and different scaling factors, are given in fig. 6.12. As we see, the points depend significantly on the value of  $A$  used for signal power manipulation. Note that according to theorem 6.4.6, for any band-limited  $b(k)$  we always have an infinite number of such points.

The procedure of adding the eigenmodes to the  $b$ -modulated signal while keeping its exact localisation, is as follows.

- (i) Modulate the waveform  $u(k)$  with the given information and according to the desired temporal support  $L$  of the signal.
- (ii) Choose the desired signal power (without eigenvalues, as in [47]) and find the appropriate scaling factor  $A$  from eq. (6.43). Further define the  $b$ -function as  $b(k) = Au(k)$ .
- (iii) For this  $b(k)$ , find point(s)  $k_{\text{eig}} \in \mathbb{C}_+$ , which satisfy  $b^*(k_{\text{eig}})b(k_{\text{eig}}) = 1$ .
- (iv) Derive corresponding  $a(k)$  via eq. (6.18), and calculate the resulting  $r(k)$  via eq. (6.19);
- (v) For each eigenvalue, calculate the corresponding  $b_{\text{eig}} := b(k_{\text{eig}})$ , which uniquely defines the respective norming constant  $c_{\text{eig}}$  via eq. (3.21).
- (vi) Generate the signal from the scattering data  $r(k)$  and set of  $\{k_{\text{eig}}, c_{\text{eig}}\}$  via any inverse NFT procedure, e.g. through the Darboux transform, given in subsection 3.3.3 or by solving the inverse problem directly with the account of discrete modes.

We perform the numerical mapping to the time domain from the scattering data using the layer-peeling algorithm (in particular, its fast implementation [142]) with the subsequent

Darboux transform, see subsection 3.3.3 that adds discrete nonlinear modes to the continuous ones without affecting the latter. The whole procedure follows the scheme given above. The results of the signal generation for both initial waveforms used for  $b$ -modulation and different scaling factor  $A$  values are given in figs. 6.13–6.14.

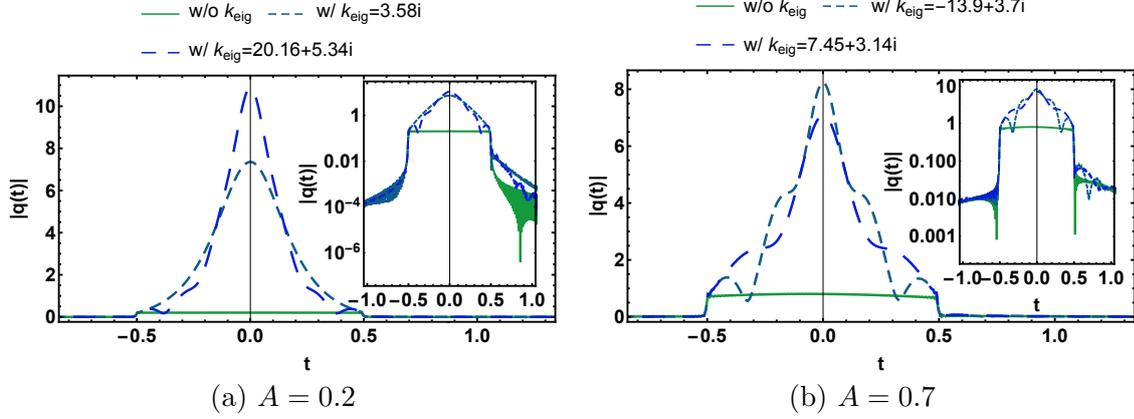


Figure 6.13: The signals, generated from the Nyquist waveform via  $b$ -modulation with (dashed) and without (solid) additional eigenvalues embedded, for different scaling factors  $A$  and different eigenvalues  $k_{\text{eig}}$ .

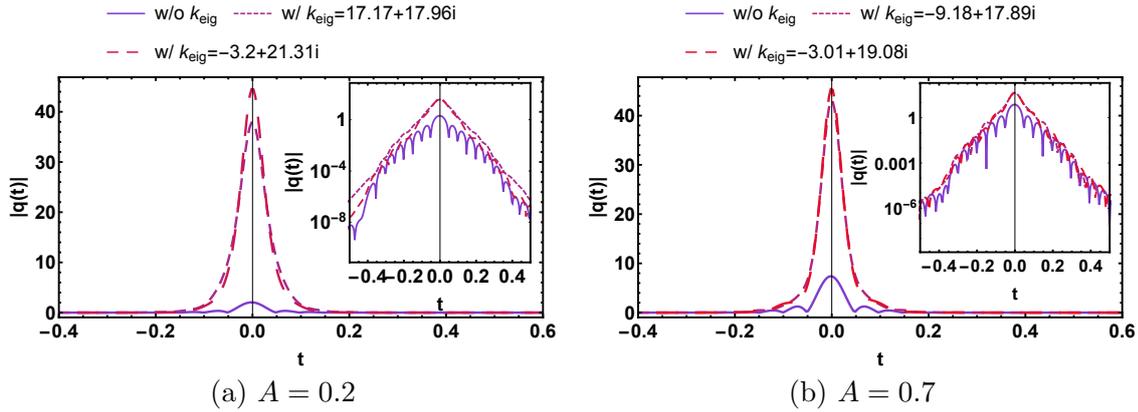


Figure 6.14: The signals, generated from the flat-top waveform via  $b$ -modulation with (dashed) and without (solid) additional eigenvalues embedded, for different scaling factors  $A$  and different eigenvalues  $k_{\text{eig}}$ .

We observe that the signals with additional solitonic eigenvalues have at least not worse localisation than the initial  $b$ -modulated signal without discrete eigenmodes, in accordance with our theory. However, the numerical algorithms introduce additional errors, which somewhat deteriorates the expected perfect localisation of the resulting  $q(t)$  profile. It can be better seen from the logarithmically scaled plots, given in the insets, that the signal tails decay rates for the profiles with and without additional discrete eigenmodes coincide almost exactly. In spite of the observed insignificant numerical errors, the results in figs. 6.13–6.14 evidently, confirm the correctness of the analytical statements presented in this work.

The addition of the eigenvalues to the  $b$ -modulated signal would be beneficial if this process provides some additional perspectives for the modulation or improvement of the transmission quality. Nonetheless, this process may also introduce some additional penalties because of the complex structure of the transmission line.

It is known that the deviations of the optical channel from the integrable NLS lead to

the effective interaction between continuous and discrete NF spectra [42], [J5, CC4, CC5, CC6] whilst in the ideal NLS model they stay decoupled. Commonly, such deviations are due to fibre loss and amplification noise, find more in subsection 2.1.4. We studied whether the presence of the bound states influences the system's response to noise in the time domain. To make the comparison clear, we introduce two separate quantities,  $\text{SNR}_b$  and  $\text{SNR}_q$ , as measures of the noise affecting the functions  $b(k)$  and  $q(\tau)$ , correspondingly. Starting from the function  $b(k)$  and optionally embedding the additional soliton, we then compute INFT to find the optical field waveform, add AWGN in the time domain, and use the direct NFT to evaluate the  $b(k)$  of the noisy signal. Then we compare the initial and back-computed  $b(k)$  to evaluate the value of  $\text{SNR}_b$ :

$$b(k) \xrightarrow{\text{INFT}} q(\tau) \xrightarrow{\text{AWGN}} q_{\text{noise}}(\tau) \xrightarrow{\text{NFT}} b_{\text{noise}}(k), \quad (6.46)$$

$$\text{SNR}_q = \frac{\int |q(\tau)|^2 d\tau}{\int |q(\tau) - q_{\text{noise}}(\tau)|^2 d\tau}, \quad \text{SNR}_b = \frac{\int |b(k)|^2 dk}{\int |b(k) - b_{\text{noise}}(k)|^2 dk}. \quad (6.47)$$

Note that to make the comparison fair, we count the signal power for the  $\text{SNR}_q$  definition as it would be *in the absence of the solitonic eigenvalue*, i.e. considering only continuous spectrum energy  $\varepsilon_{\text{cont}}$  contribution. We identify the qualitative difference depending on the signal amplitude, see fig. 6.15.

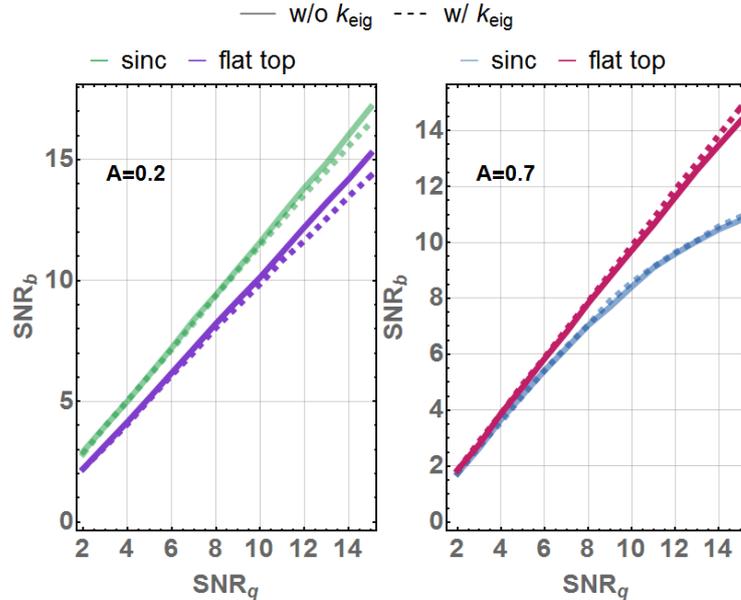


Figure 6.15: The effective SNR of the  $b(k)$  waveform, arising from the noise in the time domain, characterised by the SNR of the  $q(t)$  function, for different values of the waveform scaling factor  $A$ .

For the small-amplitude case ( $A = 0.2$ , left panel in fig. 6.15), we observe a slight deterioration of the SNR for the back-computed  $b(k)$ , if the signal contains the solitonic mode. However, we see that for the larger amplitude ( $A = 0.7$ , right panel in fig. 6.15), the value of the SNR for the soliton-free profile is smaller than that in the presence of the additional soliton mode. As far as the higher amplitudes are more relevant for NFT-based transmission where NFT can have an advantage over conventional modulation, so we can conclude from our test that the introduction of the solitonic modes may even potentially improve the overall transmission quality.

Another issue in question can be the growth of PAPR in the presence of eigenvalues (e.g. for profiles in fig. 6.13a, PAPR of the purely  $b$ -modulated waveform is 6.7 dB, whilst additions of the solitonic modes increase it up to 11.8 dB and 13.5 dB, respectively). This may cause some undesired transmitter-induced nonlinear effects. However, we would like to note that the procedure described in this paper allows us to embed eigenmodes from an infinite set. So one can pick several solitonic modes with specially chosen parameters to generate signals with the lower PAPR.

### 6.4.3 $b$ -modulation transmission systems

In this subsection, I provide the details and results of the numerical simulations of the transmission systems based on  $b$ -modulation (with added discrete spectrum modes). The results here have been published as the conference contributions [C1, C4], however, some details and discussions here have not been published.

The abovementioned limitation of the continuous spectrum energy is one of the biggest constrains limiting the signal power and, therefore, the advantage of NFT-based transmission over the conventional methods. It pushes us to seek for additional degrees of freedom which can increase the power and provide modulation opportunities. The method from subsection 6.4.2 demonstrates how to find the set of allowed bound states, which conserve the localisation. However, all discrete spectrum components are predefined by the continuous spectrum waveform, so not degree of freedom for modulation is left (besides the number of the embedded eigenvalues).

We have to admit that in practical systems the exact mathematical identities are not achievable and always violated by the imperfections, and especially by the deviations of the channel model from pure NLS. Therefore, the way of overcoming the barriers introduced by  $b$ -modulation is in changing the system regardless to the constrains for  $b(k)$ , keeping the divergence of the resulting signal out of the dedicated time window relatively small.

In the transmission systems below, I implement Nyquist-modulated  $b(k)$  waveforms, starting from the chosen constellation (see subsection 2.6.1) defined as:

$$b(k) = A \sum_{m=-(N_{\text{ch}}-1)/2}^{m=(N_{\text{ch}}-1)/2} c_m \text{sinc}(0.5kL + \pi m), \quad (6.48)$$

recalling that  $L$  is expected time-domain signal extent.

We follow the simulation approach from [86], where the expected dispersion broadening  $\kappa$  is chosen as a control parameter, using eq. (2.3).  $b$ -modulation also introduces the optimal number of signal channels determined by the transmission distance as:

$$N_{\text{ch}} = \pi B^2 L |\beta_2| / \kappa. \quad (6.49)$$

In order to get the signal with the desired bandwidth  $B$  and to satisfy the orthogonality of modes, the signal extent should be  $t_{\text{oc}} = 2N_{\text{ch}}/B$ .

Meanwhile, as the signal power in the time domain is bounded, i.e.  $A \leq 1/\max_k |b(k)|$ , the remaining signal parameters keep having a nonlinear correspondence. Fixing the (normalised) bandwidth  $B_{\text{norm}}$ , we can arrive at the maximum achievable signals amplitude:

$$A_{\text{max}} = \frac{1}{1 + 2/\pi H(\frac{N_{\text{ch}}-1}{2})}, \quad (6.50)$$

where  $H(\cdot)$  is the harmonic number function. For sufficiently large  $N_{\text{ch}}$  the sum of sinc functions can be approximated as

$$b_{\approx}(k) = 1 - \frac{2k^2}{\pi^3 B_{\text{norm}}^2}, \quad (6.51)$$

so one can arrive at the expressions connecting the bandwidth, power, and other transmission parameters:

$$\varepsilon_{\max} \approx -\frac{1}{\pi} \int_{-\pi B_{\text{norm}}/2}^{\pi B_{\text{norm}}/2} \log(1 - |b_{\approx}(k)|^2) dk \approx 3.17 B_{\text{norm}}, \quad P_{\max} \leq \frac{3.17}{4\pi} \frac{\kappa}{\gamma L}. \quad (6.52)$$

For the EDFA and propagation distance  $L = 960$  km, those expressions give the upper bound of achievable power to be around -1 dBm, but the truly achievable values of power are even lower.

Here we implement the transmission systems, based on the combination of one and several solitonic modes to reach higher data rates and to investigate the noise tolerance in the presence in the absence of the embedded eigenvalues. In these simulations, we use the solitonic phase, i.e. phase factor of spectrum amplitudes  $c_j$ , eq. (3.21), as a data carrier. We also provide a comparison of the performance of  $b$ -modulated systems with and without solitons, and conventional (OFDM) system.

The corresponding OFDM transmission is simulated for the same bandwidth and noise intensity. Insofar as the NFT approaches the conventional FT with  $\omega = -2k$  in the linear limit [135], the spectrum (6.48) corresponds to the following signal:

$$q_{\text{OFDM}}(\tau) = A \sum_{m=-(N_{\text{ch}}-1)/2}^{m=(N_{\text{ch}}-1)/2} c_m \exp\left(\frac{-2\pi i m \tau}{L}\right) \frac{2\pi}{L} \text{rect}(\tau, L/2), \quad (6.53)$$

where  $\text{rect}(x, a)$  is the rectangle with unit amplitude and localised support  $x \in [-a/2, a/2]$ .

The one solitonic state added to the signal is located at the middle of the signal extent and its shape is chosen to match the signal duration as  $A = 40/L$ . The soliton velocity and position are set to zero, so using expressions from subsection 3.1.3, the eigenvalue and the spectral amplitude are given as  $k_{\text{eig}} = iA/2$  and  $c(k_{\text{eig}}) = A_{\text{sol}} e^{i\varphi}$ , where  $\varphi$  are chosen from the 16-PSK constellation. The full schematics can be found in fig. 6.16

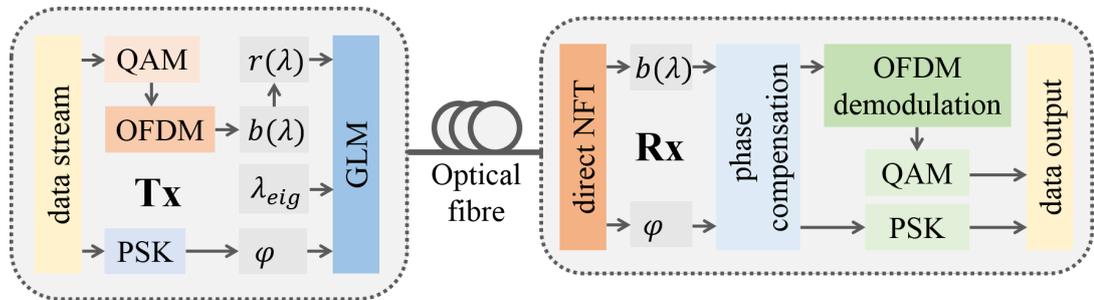


Figure 6.16: Schematics of  $b$ -modulation transmission system with one additional solitonic mode modulated.

The transmission of the optical signal in SMF with EDFA amplification over 80 km span is simulated for propagation distances  $Z = 480$  km and  $Z = 960$  km. The signal bandwidth is  $B = 40$  GHz. The simulation results are presented in fig. 6.17. In the noiseless transmission, fig. 6.17a, as it is expected, we observe the better performance of  $b$ -modulation system compared to the OFDM format. We also see that the presence of discrete eigenvalue effectively reduces the Q-factor of transmission, and this reduction is more noticeable for longer fibre links. The Q-factor degradation for high signal powers arises from the numerical error of GLM algorithm used in the INFT block.

However, in the system with optical amplification, see fig. 6.17b, the noise becomes the main source of error and the proposed combined modulation leads to a slight increase of Q-factor (approximately on 0.5 dB). Together with this, we have the 1 Gbps increase

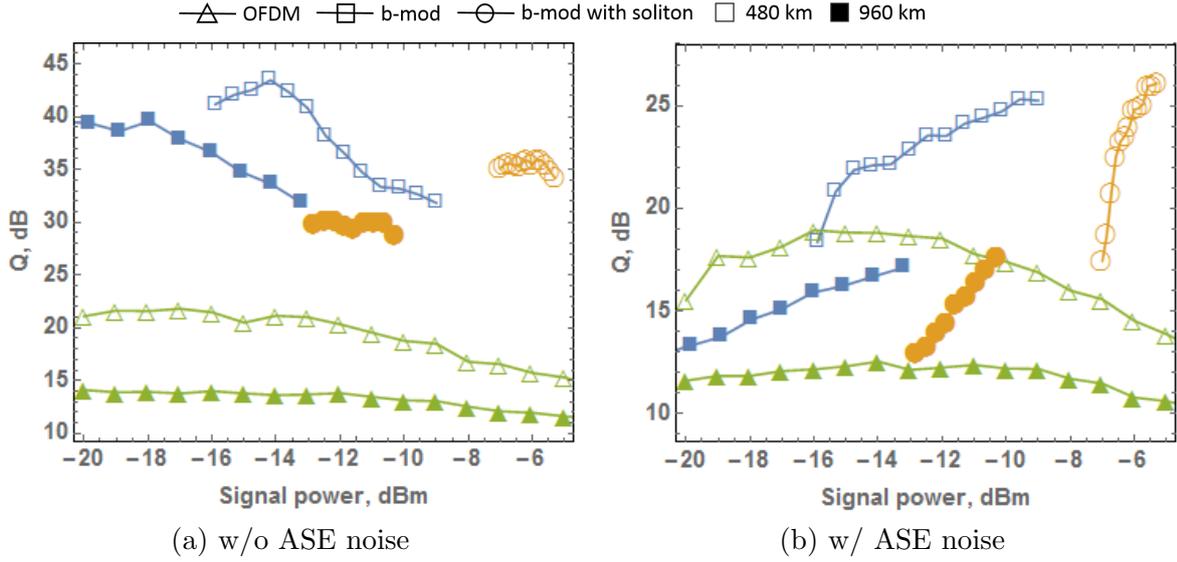


Figure 6.17:  $Q$ -factor dependency on signal power in  $b$ -modulated systems with and without embedded (single) eigenvalue, compared with corresponding OFDM transmission.

in data rate for the 480 km link length. The obtained results clearly indicate the good potential of this kind of systems to try several solitonic modes.

For several solitons, it is important as well to locate them within the signal temporal window, but now we cannot simply ignore the position and velocity for eigenvalues separation. Choosing the set of eigenvalue  $k_j$ ,  $j = 1 \dots N_s$ ,  $N_s = 7$ , define the spectral amplitude via  $b_j$  as:

$$b(k_j) = \exp(A_{\text{eig}} T_j) \exp(-i\varphi_j), \quad j = 1, \dots, 7, \quad (6.54)$$

and since  $\Re k_j$  is responsible for discrete modes drift during the propagation along the fibre, we use symmetric locations within the time window  $T_j = (j - (n + 1)/1)\Delta$ . It is worth noting, that because  $b(k)$  exponentially increases in the upper complex half-plane, the correct computation of  $b(k_j)$  for complex  $k_j$  requires a forward-backward approach to the solution of (3.5), see [10], [J2].

In fig. 6.18, the scheme of the communication system with stages of the processing on the Tx and Rx is presented.

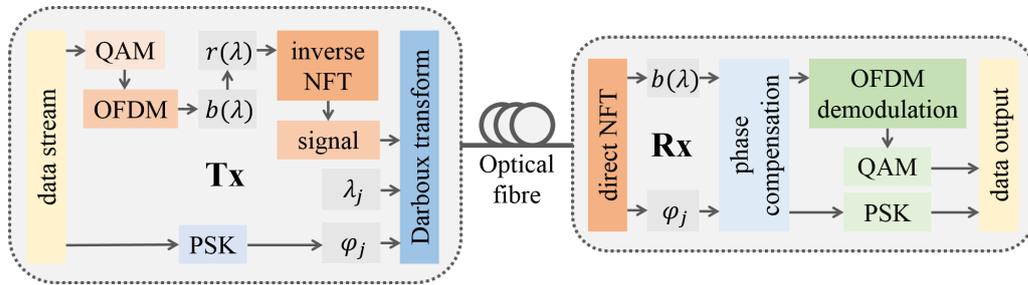


Figure 6.18: Schematics of  $b$ -modulation transmission system with seven additional solitonic modes modulated.

The transmission of the optical signal in SMF with EDFA amplification over 80 km span is simulated for propagation distance  $Z=480$  km. The signal bandwidth is  $B = 30$  GHz. We use 64-QAM constellation for continuous spectrum coefficients in (6.48) and 32-PSK constellation in (6.54). The simulation results are presented in fig. 6.19.

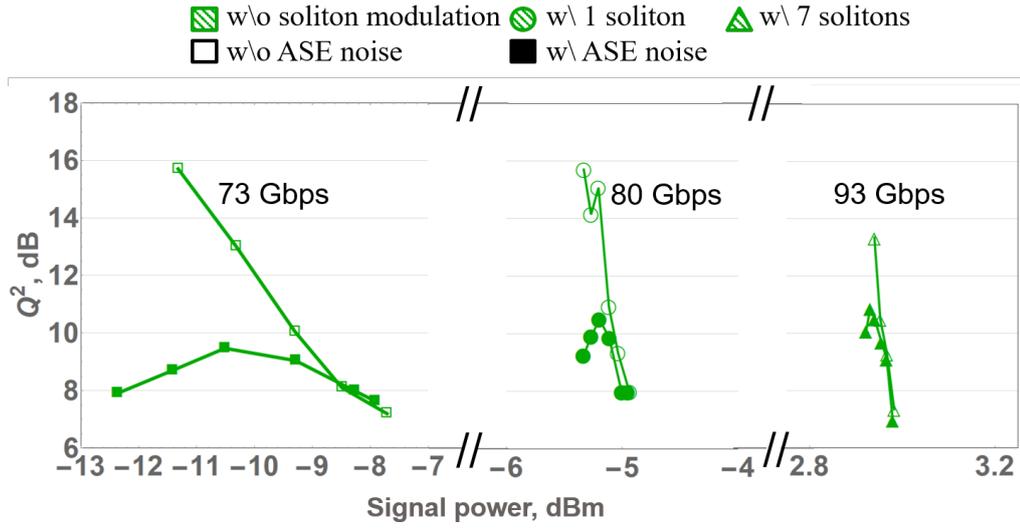


Figure 6.19:  $Q$ -factor dependency on signal power in  $b$ -modulated systems with and without embedded eigenvalues.

As we can see from the simulations, the additional discrete eigenvalue modes allowed to increase the profit in the achievable data rate (see labels in the fig. 6.19) from 10% to 27%. Moreover, due to the increase of the effective SNR, we observe a small gain in terms of systems'  $Q$ -factor in the case of discrete spectrum modulation. With regards to the processing complexity, we inevitably get some penalty in runtimes when adding the discrete modes. The Rx runtimes are less affected: 0.5% penalty for one discrete eigenvalue and 2.5% penalty for 7 eigenvalues. On the Tx, we observe larger increase: 4% for 1 additional discrete state and 20% for seven additional states. However, we must notice that for more discrete modes the runtime increase is smaller than the gain in terms of data rate (factor 5 runtime penalty comparably to factor 7 data rate gain).

## 6.5 Periodic boundary conditions

The concept of NFT-based communication can be implemented for NLS model with both vanishing boundary conditions, section 3.1, and periodic boundary conditions, section 3.2. Both models can be applied to some extent, however, for these cases, the picture in NF domain differs principally, with various degrees of freedom to modulate the information in. The periodic case is less popular among the community, both from the mathematical side, and communicational side. To the best of our knowledge, the general problem of parametrisation of arbitrary exactly periodic NLS solution by means of IST is not solved. However, the conventional NFT approach in communication run out of its resources and opportunities, which motivates to redirect research in this area to some alternative route. The way of utilizing the periodically continued NFT-generated profile is known in the literature as PNFT.

For the illustration of the following motivation part, refer to fig. 6.20.

In practical systems, the assumption of the infinite line (i.e.  $q(\tau)$ ,  $\tau \in (-\infty, \infty)$ ) is not realistic, and therefore, the truncated waveforms are used. When the “insufficient portion” of the signal is processed, or when the neighbouring NFT symbols overlap due to a weak time-decay rate or dispersive spreading, the noticeable performance penalty emerges [25, 26]. Also, aside from some technical problems related to the utilisation of “unusual” NFT-generated pulses (the systems' sensitivity to transceiver imperfections [80, 135] and optical noise [26]), and the issues related to the NFT routines performance

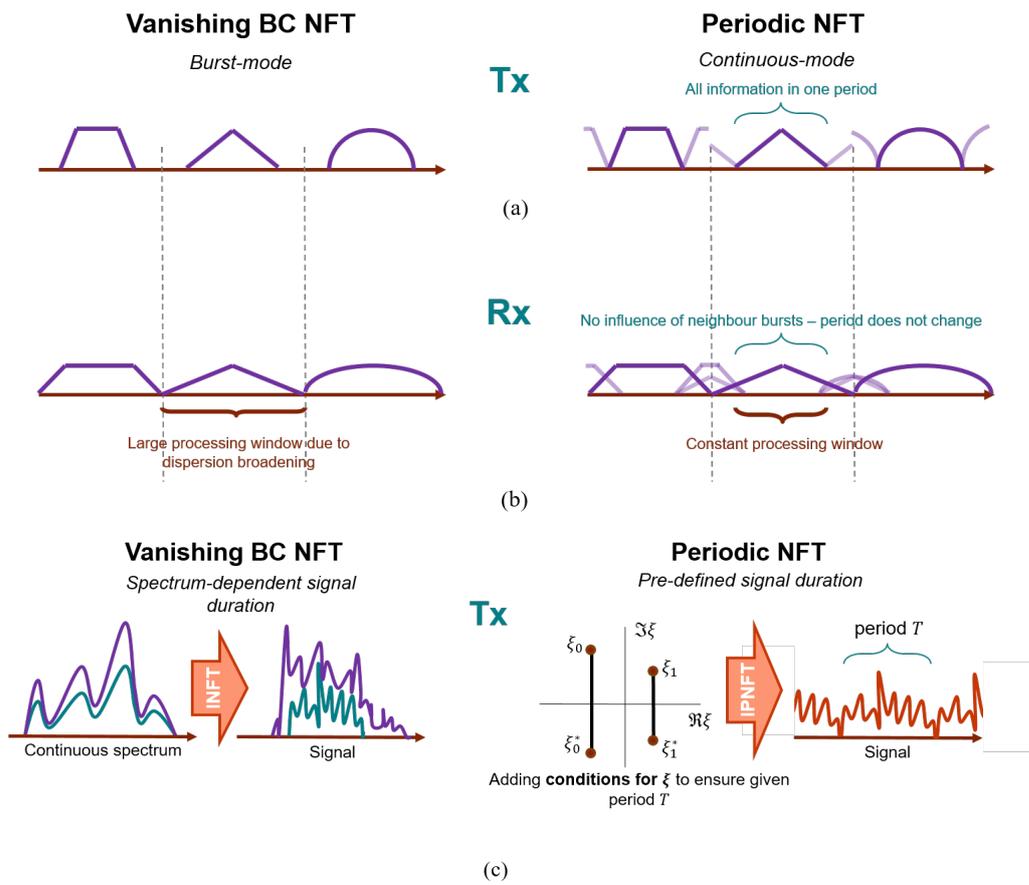


Figure 6.20: The illustration of the advantages of PNFT comparably to the (conventional) NFT from an implementation perspective.

as the signal power and/or processing interval grow [25, 84, 89, 151], which are rather just the peculiarities of the method, we mention more fundamental issues: (i) the hardly-controllable (and typically extended) optical signal duration after the inverse NFT at the Tx [80, 81, 117]; (ii) a large processing window required, involving the full dispersion-induced memory; (iii) high PAPR values [135].

In this respect, considering the *periodically-continued* signals, one can relax or even eliminate some of the afore-stated problems [63]. In particular, the period of PNFT-generated signal and, hence, the symbol duration are well controllable parameters. The PAPR of periodically-continued signals is much lower due to the absence of zero-padding “wings” in-between the information-bearing time intervals (symbols). Within the PNFT one has to process just one (predefined) period reducing the processing window and, hence, the number of samples to process [63], and the overall processing complexity [22]. The SNR degradation due to signal-noise interference, proportional to the processing interval duration for the transmission inside the NFT domain [26, 32], can be diminished due to the shorter processing window as well. A PNFT-based system can have, in addition, a better overall SE due to the smaller time-bandwidth product of the signal [64]. Thus, PNFT applied to optical transmission allows us to retain the “conventional” NFT advantages, but can render a lot of beneficial features itemised above.

Within the PNFT-based communication concept, we have to deal with the periodic solutions of nonlinear integrable equations. The latter has been the subject of study for a long time [17, 58, 76, 96, 152], and their vast application area ranging from water gravity waves to optical turbulence, has encouraged researchers to develop a solid mathematical base for their description [5, 21, 110].

The main challenge in PNFT-based communication arises from the inverse transformation at the Tx side: constructing a periodic signal in the time domain given the modulated NFT spectrum components. To the best of our knowledge, there is still a lack of a generic and effective method to perform the inverse PNFT, and more active involvement of mathematicians would be greatly beneficial here. There exist approaches suitable for particular nonlinear solutions and respective PNFT spectra types [130], however, these methods have somewhat limited flexibility compared to quite versatile conventional NFT transmission approaches.

In this section, I provide results of the research where the periodic signal is created by means of the suitably designed RHP, and the corresponding NF image is used for control of signal parameters and information modulation. The work here is done in collaboration with Dr Kamalian and Prof Shepelsky. It is published in the form of two papers [J1,SJ1], and some results were delivered during the topical conferences [C2, C6, C7].

### 6.5.1 Signal properties for the RHP-generated profiles

In this subsection, I provide the dependence between the signal parameters, namely, power  $P$ , bandwidth  $B$  and period  $T$ , defined by the structure of the main spectrum of the periodic NLS solution. The research similar to this, I have performed during the design of the communication systems for [J1] and [SJ1], however, here I present a more complete and general study.

Let us start with genus- $\mathcal{N}$  solution of NLS, which is directly determined by  $\mathcal{N} + 1$  main spectrum points. The structure of the main spectrum, looking as slightly horizontally stretched uniform symmetric grid, see fig. 6.21, reduce the number of parameters to only three:  $\mathcal{N}$ , the spacing between the cuts  $\Delta\Re\lambda$ , and (common) imaginary part  $\Im\lambda$ . During this research, we also observed that the corresponding phases do not give a noticeable contribution to the signal parameters, see, e.g., in fig. 6.22, the normalised variation of the signal power for constant main spectrum and random (uniformly distributed) phases.

With regards to the signal period, the zero-order approximation of the system (3.92)

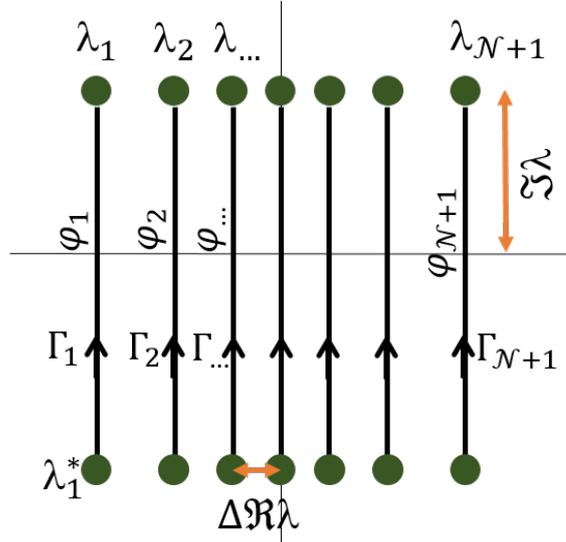


Figure 6.21: The structure of the NF spectrum of finite-band solution, adjusted to be periodic, with the main spectrum points  $\lambda_j$ , contours  $\Gamma_j$  and the associated phases  $\phi_j$  depicted.

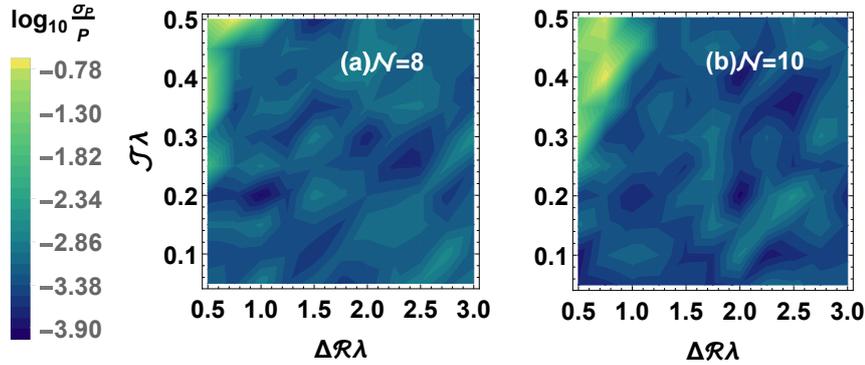


Figure 6.22: Normalised power deviation  $\sigma_P/P$  as a function of the main spectrum structure, defined by  $\Delta\Re\lambda$  and  $\Im\lambda$ , for high-genus periodic RHP profiles.

returns the period value as  $T_{\text{norm}} = \pi/\Delta\Re\lambda$ . As for the signal bandwidth, it is given as  $B_{\text{norm}} = 2\pi\mathcal{N}\Delta\Re\lambda$ .

We can further develop the next order of approximation for the link between the main spectrum structure and signal parameters. Our goal is to be able to manipulate power, keeping data rate and the design constant. It means, we are limited by the constraints to maintain the number of bands  $\mathcal{N}$ , bandwidth  $B_{\text{norm}}$  and duration  $T_{\text{norm}}$ . While the level curves for a constant period are straight lines in  $(\mathcal{N}, \Delta\Re\lambda)$  space, more complicated level curves of constant bandwidth  $B_{\text{norm}}$  are presented in fig. 6.23.

From the numerical simulations, where we had to adjust the main spectrum to change the power only. The system is controlled by three parameters,  $\mathcal{N}$ ,  $\Delta\Re\lambda$  and  $\Im\lambda$ . Increasing power by increasing  $\Im\lambda$ , we get that the value of  $\Delta\Re\lambda$  no longer satisfies the zeroth-order approximation  $\pi/T_{\text{norm}}$ . We denote this deviation as  $\delta\Delta\Re\lambda \equiv \Delta\Re\lambda - \pi/T_{\text{norm}}$ . numerical simulations provide us a three-argument dependency  $\delta\Delta\Re\lambda(\mathcal{N}, T_{\text{norm}}, \Im\lambda)$ .

First, we assume, that all influencing parameters (those, which are maintained constant, first of all,  $T$  and  $\mathcal{N}$ ), entry the dependency in a factorised way as:

$$\delta\Delta\Re\lambda(\mathcal{N}, T_{\text{norm}}, \Im\lambda) = f(\Im\lambda) \times g(\mathcal{N}, T_{\text{norm}}). \quad (6.55)$$

From the power dependency and zeroth-order approximation, we can conclude that

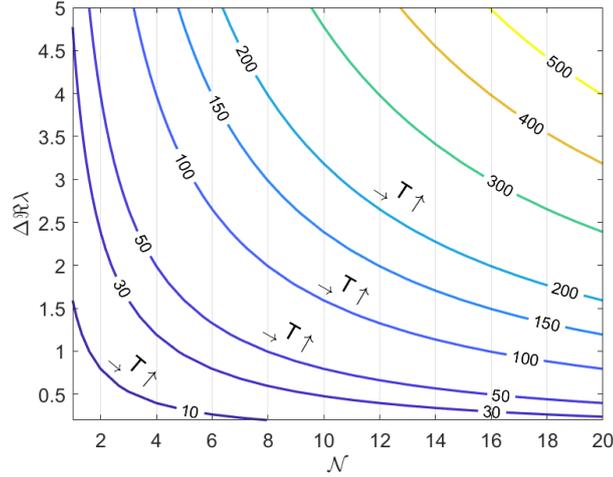


Figure 6.23: Isolines of constant bandwidth  $B_{\text{norm}}$  in zeroth-order approximation in  $(\mathcal{N}, \Delta \Re \lambda)$  plane (labels on the curves designate the normalise bandwidth value).

$f(x)$  is increasing function of a positive argument, starting from the origin, so  $f'(x) > 0, \forall x$ , and  $f(0) = 0$ . From there we can assume that  $f(x)$  is power function:  $f(x) = x^\alpha$ . From the numerical fitting, we identify the most accurate guess for a power parameter  $\alpha = 1.75$ . The further two-parametric fit (see fig. 6.24) provides a polynomial approximation, and the quadratic polynomial, sufficient for this purpose, is given by:

$$g(\mathcal{N}, T_{\text{norm}}) = -0.1446 + 0.03838\mathcal{N} - 0.009975T_{\text{norm}} - 0.00227\mathcal{N}^2 + 0.004613T_{\text{norm}}\mathcal{N} + 0.004151T_{\text{norm}}^2. \quad (6.56)$$

Finally, in 3D  $(\mathcal{N}, \Delta \Re \lambda, \Im \lambda)$  space, the necessity to maintain data rate leads to the following traces for changing parameters to increase power only, see fig. 6.25.

### 6.5.2 Numerical accuracy of RHP-based processing

The most demanding part of the inverse transformation stage to get a time-domain profile, is to solve the RHP. The computational complexity of calculating the value of the signal at each point in time and distance via the numerical RHP solution is of the order  $O(N_\lambda \log N_\lambda)$ , where  $N_\lambda$  is the total number of points used to represent the otherwise continuous cuts in the complex  $\lambda$ -plane. Here we assume that these discretization points are uniformly distributed over these cuts. We note that when considering the specific structure of the cuts, the uniform grid might not necessarily be the best option, and some more efficient strategies for the RHP cuts discretization may exist. In fact, our numerical simulations show a marginal impact of changing  $N_\lambda$  on the accuracy of calculations after a certain limiting  $\Delta \Im \lambda$  value, the step between the discretization points in our cuts. On the other hand, increasing the number of temporal samples, i.e. diminishing the sample value  $\Delta \tau$ , have a greater impact on the numerical accuracy of the direct and inverse transformations and on the overall performance of the resulting communication system. The quality of the back-to-back phase computation for our recovering the values of randomly modulated phases  $\phi_j$  is shown in fig. 6.26 in terms of the  $Q^2$ -factor density plot vs  $\Delta \Im \lambda$  and  $\Delta \tau$ . The back-to-back procedure consists of the inverse and direct transformations and accounts for the combined errors produced by both of them together.

Signal power is directly influenced by the imaginary part of the main spectrum and almost independent of the auxiliary spectrum (see fig. 6.22).

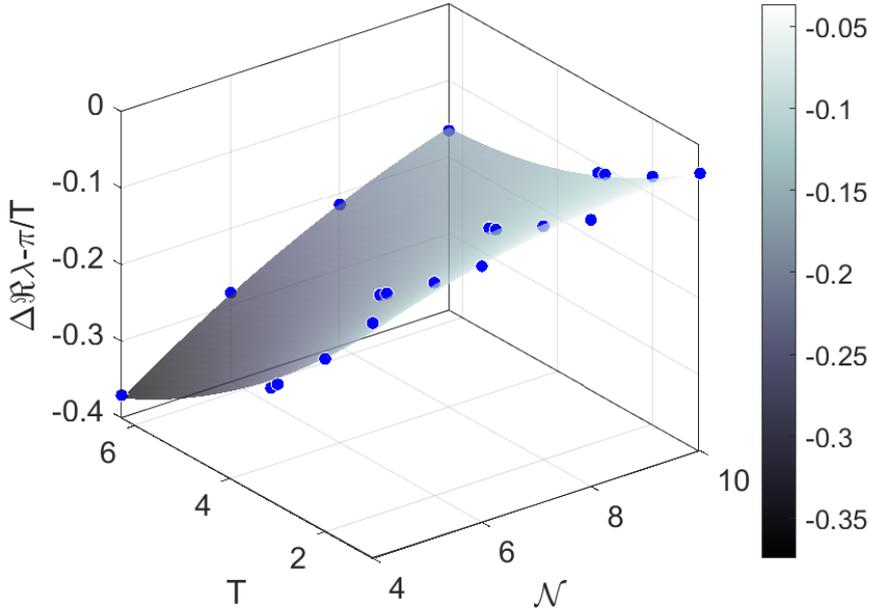


Figure 6.24: Two-parametric fit of a deviation of main spectrum point spacing  $\delta\Delta\Re\lambda$  as a function of signal period  $T_{\text{norm}}$  and the number of bands  $\mathcal{N}$ . From this plot we conclude that quadratic fit is sufficient for a factor function  $g(\mathcal{N}, T_{\text{norm}})$  (goodness of fit  $R^2 = 0.98$ ).

We generally mention two limitations for increasing the signal power up to an arbitrary level. One is the numerical error ensued from the direct transform stage (see [141, 90]) and in the inverse transform stage by deteriorating the spectral resolution,  $\Delta\Im\lambda$  (see fig. 6.26). To explain the latter, note that as signal power increases by raising the imaginary part of the main spectrum, a fixed number of points representing the cuts leads to larger  $\Delta\Im\lambda$ , which in turn, gives rise to less accurate numerical results. The other constraint is related to our approach of finding the appropriate main spectrum to guarantee the periodicity of the signal.

As for the signal bandwidth, since increasing the bandwidth (while keeping the signal duration fixed) is done by increasing the number of cuts in the main spectrum, finding appropriate main spectrum entails solving a system of nonlinear equations with more unknown values. This, in turn, makes solving eq. (3.92) numerically more difficult in the general case.

### 6.5.3 Main spectrum modulation in RHP-based transmission

In this part of the research, we ignore the existence of associated phases and use the main spectrum location as an information carrier. As a proof-of-concept, this system is based on the simplest non-trivial case, i.e., the case  $\mathcal{N} = 1$ . In this case, the problem of ensuring periodicity is greatly simplified: it is sufficient to ensure that  $f_0 = 0$  in (3.94), which, in turn, can be done by simultaneously shifting all the spectral data points  $\{\lambda_j\}_{j=0}^{\mathcal{N}}$  along the real axis. Thus a preliminary step in the construction of a genus-1 solution consists in the adjustment of the main spectrum: given  $\{\hat{\lambda}_j\}_{j=0}^1$ , the real-valued shift is computed as follows:

$$\hat{f}_0 = \frac{\int_{\hat{\Gamma}_1} \frac{\xi d\xi}{\hat{w}(\xi)}}{\int_{\hat{\Gamma}_1} \frac{d\xi}{\hat{w}(\xi)}} - \frac{1}{2} \sum_{j=0}^1 (\hat{\lambda}_j + \hat{\lambda}_j^*) \quad (6.57)$$

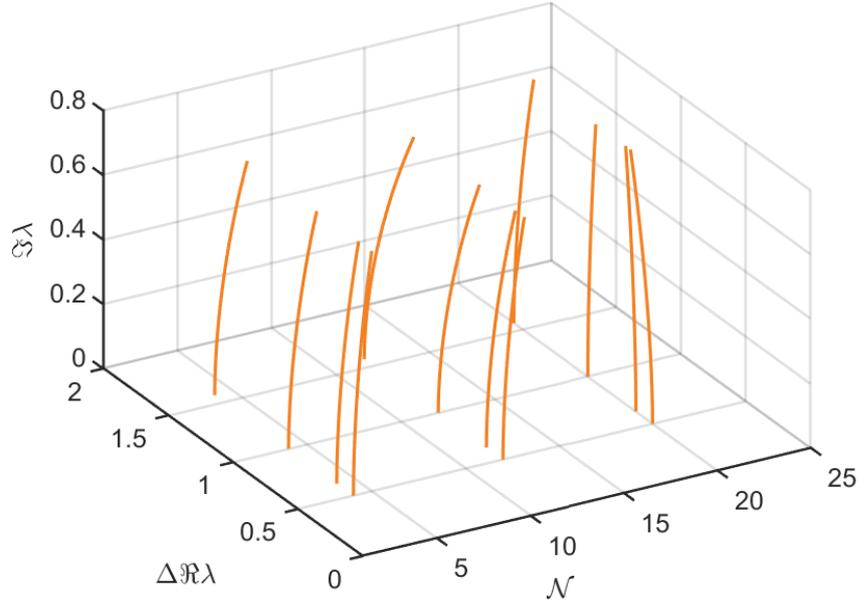


Figure 6.25: Curves for power manipulation maintaining data rate in space of main spectrum structure parameters  $(\mathcal{N}, \Delta\Re\lambda, \Im\lambda)$ .

with  $\hat{w}(\xi) = \hat{w}_+(\xi)$  defined as eq. (3.44) from the preliminary set of the main spectrum points  $\{\hat{\lambda}_j\}_{j=0}^1$ . Then the new points  $\lambda_j$ , corresponding to a periodic genus-1 solution with  $f_0 = 0$  (see (3.94)), are defined by:  $\lambda_j = \hat{\lambda}_j + \hat{f}_0$ .

Now, given  $\{\lambda_j\}_{j=0}^1$  and  $\phi_1$  (setting  $\phi_0 = 0$ ), the procedure of construction of a genus-1 solution is reduced to the following steps:

(i) Calculate

$$C^f = -\frac{2\pi i}{\int_{\Gamma_1} \frac{d\xi}{w(\xi)}}, \quad C^g = C^f \sum_{j=0}^1 (\lambda_j + \lambda_j^*), \quad (6.58)$$

with  $\hat{\lambda}_{0,1}$  replaced by  $\lambda_{0,1}$ .

(ii) Define the jump matrices:

$$\mathbf{G}_0 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}, \quad \mathbf{G}_1(\tau) = \begin{bmatrix} 0 & -1/\vartheta(\tau) \\ \vartheta(\tau) & 0 \end{bmatrix}, \quad (6.59)$$

where  $\vartheta(\tau) = ie^{i(C^f\tau + \phi_1)}$  and  $C^{f,g}$  are defined by (6.58).

(iii) Solve the RHP: given  $\mathbf{G}_j$ ,  $j = 0, 1$ , find a  $2 \times 2$ -valued function  $\mathbf{M}(\cdot, \cdot, \lambda)$  analytic in  $\mathbb{C} \setminus \{\Gamma_0 \cup \Gamma_1\}$  and satisfying the following conditions:

$$\mathbf{M}^-(\tau, \lambda) = \mathbf{M}^+(\tau, \lambda) \mathbf{G}_j(\tau), \quad \text{for } \lambda \in \Gamma_j, \quad j = 0, 1, \quad (6.60)$$

$$\mathbf{M}(\tau, \lambda \rightarrow \infty) \rightarrow \mathbf{I}.$$

(iv) Retrieve the solution of NLS by eq. (3.81).

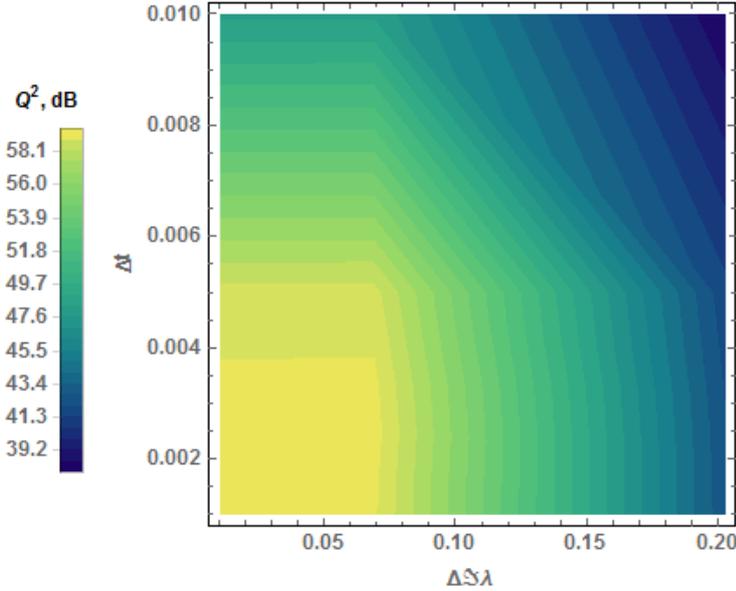


Figure 6.26: The  $Q_{\text{EVM}}^2$ -factor of the B2B calculation of phases  $\phi_j$ , embedded into the signal. The signal is generated via the RHP solution for the 15 bands.  $Q^2$  is computed as the function of the cut sampling step  $\Delta S\lambda$  for the vertical cuts, and time sampling step  $\Delta\tau$ .

We refer the reader to detailed book [134] for the numerical realisation of this step. For our purpose, in this work we will use the existing numerical package `RHPackage` [107]. Using `RHSolve` subroutine from `RHPackage` we note that in order to solve the RHP we need to deal with fast discrete cosine transform (DCT) and compute Cauchy integral, see eq. (3.95) at  $N_\lambda$  Chebyshev points of the second kind, where  $N_\lambda$  is the total number of spectral points on the arcs (directly related to the resolution  $\Delta\lambda$  between the adjacent discretisation points). The numerical error is shown to decay spectrally as  $N_\lambda$  grows [109]. Computing (3.95) can be significantly expedited by expanding the solution using the Chebyshev polynomials of the first kind in as much as the expressions for the Cauchy integral involving these polynomials are known explicitly. Evaluating the  $N_\lambda$  Chebyshev polynomials can be done using  $O(N_\lambda \log N_\lambda)$  floating-point operations, which along with the same figure for the DCT makes the total computational complexity of solving the RHP for each time sample of the solution  $O(N_\lambda \log N_\lambda)$  [109]. In this respect, we note that increasing the genus, which is equivalent to increasing the number of arcs participating in the RHP, does not increase the complexity in the numerical implementation of the RHP approach. This feature is the principle advantage with respect to the alternative approach of the construction of algebro-geometric solutions where the Riemann theta functions are used. Note also that there is the more advanced version of the numerical package for RHP by the same author, written for Julia language [108], however, we did not get enough resources to use it properly for the simulations, despite its claimed high computational power.

To have a signal capable of carrying one complex QAM symbol in its main spectrum, we use a genus-1 solution as the most straightforward non-trivial case. The exemplary main spectrum is shown in fig. 6.27. For simplicity, we also do not make use of the additional jump matrices parameters  $\phi_{0,1}$  from (3.79) for our building up the simple proof-of-concept communication system here, setting  $\phi_0 = \phi_1 = 0$ .

To control the signal parameters, we recall that the difference between the real parts of  $\lambda_0$  and  $\lambda_1$  is related to the value of the period. Thus, if we want to have some predefined period, we can adjust the real parts of  $\lambda_0$  and  $\lambda_1$ , and the QAM symbol can be constructed

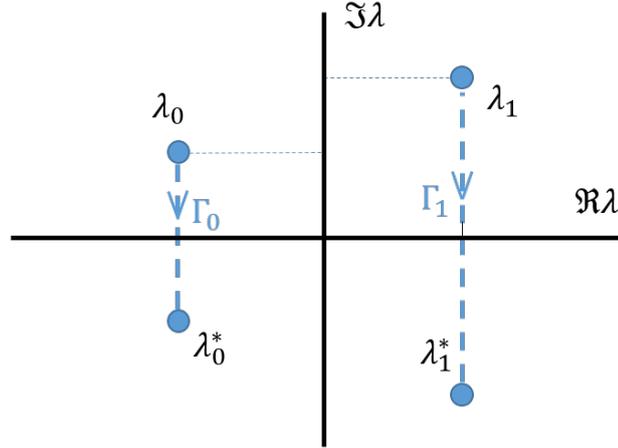


Figure 6.27: The exemplary main spectrum of a genus-1 signal: the figure shows complex  $\lambda$ -plane with the cuts  $\Gamma_j$  along the arcs  $(\lambda_j, \lambda_j^*)$ , and end-points  $\{\lambda_j, \lambda_j^*\}$ . In our proof-of-concept study the data are mapped on the imaginary parts of eigenvalues.

by modulating their imaginary parts. In other words, knowing the imaginary part of the eigenvalues and a given signal period,  $T$  (and thus fixing  $C^f = 2\pi/T$ , see (6.58)), we can find the real parts of  $\lambda_0$  and  $\lambda_1$ . The coefficient  $C^g$  in eq. (6.58) is then calculated, the jump matrices from (6.59) are defined, and the solution of the RHP (6.60) is obtained using the `RHPackage` [107]. Two example signals obtained in such a way (together with their main spectrum) are shown in fig. 6.28.

Due to the perturbing noise (numerical, receiver and amplifier noise), the signal power has a lower bound below which it is impossible to discern the information, bearing eigenvalues from the noise-induced spurious ones. On the other hand, the upper bound on the signal power is determined by the numerical errors in nonlinear signal processing and more pronounced signal-noise interaction. In the case considered in our work, a QAM constellation, made up by the imaginary parts of two main spectrum points, contains a point corresponding to a signal with the lowest power (the left bottom point of the constellation, see fig. 6.31 for example) and other points corresponding to higher signal powers (having larger  $\Im\lambda_{0,1}$ ). Increasing the signal power by raising the imaginary part of the main spectrum points makes the distance between the constellation points grow farther. Thus, the probability of the constellation points crossing the decision boundaries at the Rx gets lower. The consequence of controlling the signal power by means of changing the imaginary part of the eigenvalues (changing the SNR) and the ability to control the error by scaling the constellation is the main reason why we have chosen the imaginary parts of the spectrum points to carry the data. The real parts are determined by the desired signal period and the chosen imaginary parts. However, we note that the effective SNR defined inside the NFT domain is an involved function of the signal power and distance between the constellation points: we can not have a better performance by simply enlarging the constellation size. So the interrelation between the system performance and signal power (constellation size) is studied through numerical simulations in the following section.

The simulations are done for SMF fibre link with ideal Raman amplification and in a link with lumped amplification (using EDFA) are presented. Using such a model, we apply the adjustment explained in [64] to optimise the location of amplifiers to improve the performance of a NFT-based communication system. To overcome the ISI caused by the chromatic dispersion-induced signal broadening, we append the signal with cyclic extension in the time domain, fig. 6.29. Each signal, carrying one QAM symbol, is extended for the value greater than the channel memory calculated from the signal bandwidth, see

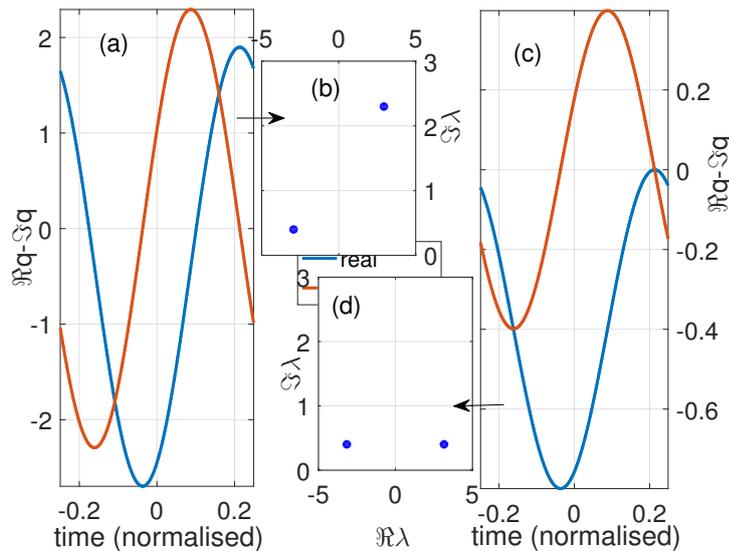


Figure 6.28: Two examples of an optical signal waveform, panes (a) and (c), constructed from the main spectra shown in panes (b) and (d), respectively (only the region  $\Im\lambda > 0$  is depicted).

subsection 2.1.2.

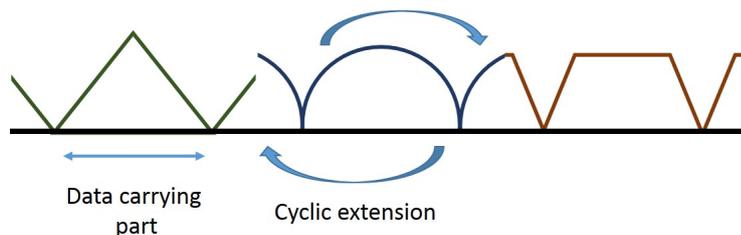


Figure 6.29: Schematic of the cyclic extension of the periodic signal within the window of the dispersion-caused signal broadening.

In our simulations, the symbols are transmitted in the bursts of  $4M$  samples, where  $M$  is the size of the constellation. The burst is formed as follows: several signals, modulated using random data, are cyclically extended and put together. At the Rx, a simple phase rotation of the constellation is used. Fig. 6.30c depicts the  $Q^2$ -factor calculated from the EVM for four systems with different constellation sizes in a 880 km link. The  $Q^2$ -factor is averaged over the number of symbols in each burst and over  $2^8$  runs. The symbol rate is 0.8 GSym/s; it can be increased by increasing the signal bandwidth (changing the real part of the eigenvalues). Fig. 6.30 shows that there is an optimum power at which the  $Q^2$ -factor is maximal. One reason for a decline in the system performance at higher powers is the dependence of numerical accuracy on power. The received constellations at optimum power are shown in fig. 6.30(a)-(b). As explained before, the axes of these scatter plots are the imaginary parts of the two eigenvalues in the discrete (main) spectrum of the received signal,  $\Im\lambda_0$  and  $\Im\lambda_1$ .

Fig. 6.31c portrays the dependency of the  $Q^2$ -factor on the link length and the impact of the particular amplification type: ideal Raman and EDFA. The close-to-ideal Raman amplification can be realised with a reasonable degree of accuracy through the second-order Raman pump.

In our simulations, the signal power was set to  $P = -5$  dBm, and the symbol rate

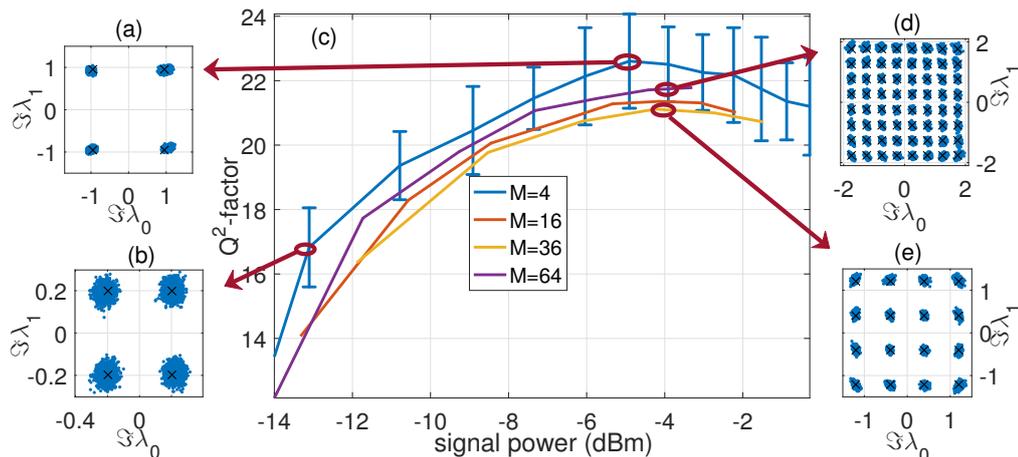


Figure 6.30: The simulation results of the transmission systems based on genus-1 periodic RHP solution: (a) the 4-QAM receiver constellation at the optimum power; (b) the 4-QAM receiver constellation at power  $P = -13.1$  dBm, (c)  $Q^2$ -factor as a function of signal power for  $Z = 880$  km link with EDFA (0.8 GSym/s symbol rate); (d) the 64-QAM, and (e) the 16-QAM receiver constellation at the optimum power.

is 1 GSym/s. The received constellation is depicted at distances  $Z = 580$  km and  $Z = 1120$  km. Another important object in communication systems' design is the probability distribution function (PDF) of the received symbols. Finding the PDF of the received symbols given the transmitted ones is necessary to find the mutual information, channel capacity and to design an optimum coding and detection strategies. Since there is still a lack of a mathematical understanding of the behaviour of the PNFT spectrum quantities under the influence of optical noise [63], here we rely on the empirical PDF coming from the histogram of the received symbols. For a 1024-QAM signal, a 2D histogram of the received QAM symbols is plotted in fig. 6.32 by using  $2^9$  symbols. The almost circular shapes in that figure indicate that the received distribution is close to the circular Gaussian one. However, as it typically occurs in other communication systems based on the NFT [32], the characteristics of this Gaussian PDF differ for symbols with different signal power. This can be seen in figs. 6.32(b)-(c) where the histogram of the small- and large-power parts of the constellation are shown.

Considering only two arcs (the simplest non-trivial case) simplifies the calculations and algorithms, however at the expense of our having a small number of the degrees of freedom per signal that can be simultaneously modulated. The lack of direct precise control over the signal bandwidth and a relatively small SE emerging due to the large time-bandwidth product is among the consequences of this choice. In a two-arc system, each signal carries just one QAM symbol. Therefore, the cyclic prefixes inserted to mitigate inter-symbol interference and carrying no data produce a negative impact on the SE value. However, we note that the important feature of our current approach is that it can be scaled up in a straightforward manner. So we can generally incorporate more arcs for our having more available parameters for the modulation, and then adjust the system parameters to have a periodic solution. Then we can employ some parameters to set the bandwidth and power.

#### 6.5.4 Phases modulation in RHP-based transmission

In order to get better over the signal parameters and increase the achievable data rate, we get back the associated phases to use them as an information carrier and consider the main spectrum, in the light of the results from subsection 6.5.1, for manipulation power, bandwidth and period. In this subsection, we demonstrate the performance of such a

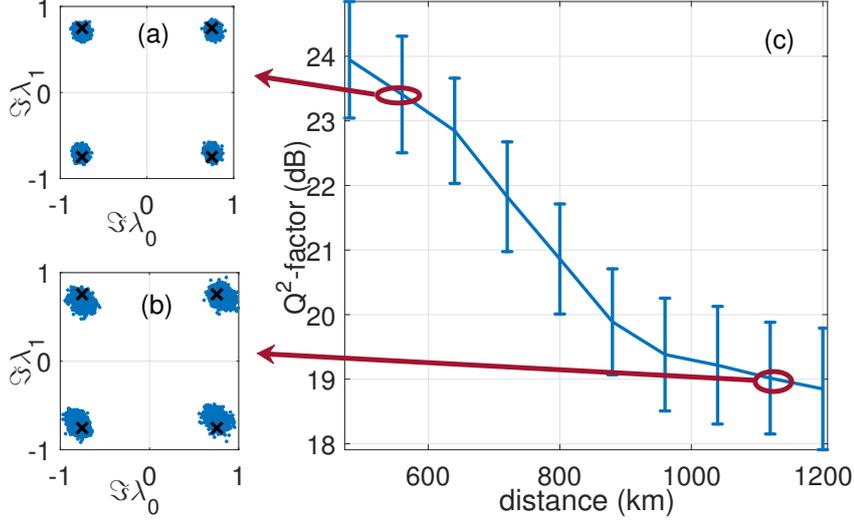


Figure 6.31: The simulation results of the transmission systems based on genus-1 periodic RHP solution: (a) the received constellation at  $Z = 580$  km; (b) the received constellation at  $Z = 1120$  km; (c)  $Q^2$ -factor as a function of the link length ( $P = -5$  dBm, 1 GSym/s symbol rate).

transmission setup.

In our system, we map the MPSK symbols ( $M = 8$ ) to the phases of a 15-band NLS solution. The resulting time-domain signal is sent to a fibre link containing 13 80-km-long spans, the total transmission distance being  $Z = 1040$  km.

For higher-genus NLS solutions, it is sufficient to require all frequencies  $\{C_0^f, \dots, C_N^f, f_0\}$ , determined uniquely by  $\lambda_j$  through eqs. (3.87) and (3.94), to be commensurable. Particularly, this holds in the case with  $C_j^f = j - N/2 - 1$  and  $f_0 = 0$ . Then, (3.87) complemented by (3.94) can be viewed as a system of  $N + 1$  complex nonlinear equations with respect to  $\lambda_j$ : given  $\{C_j^f\}_{j=0}^N$ , determine  $N + 1$  complex parameters  $\lambda_j$  entering the coefficients of that system. In order to get the periodic solution, we employ the iterative approach starting from infinitesimal  $\Im\lambda_j$ .

Let the main be a grid of equidistant points in the real direction having the same imaginary parts:

$$\Lambda = (-\lambda_{N/2}, \dots, -\lambda_1, \lambda_0 = 0, \lambda_1 = \Delta\lambda, \dots, \lambda_{N/2}). \quad (6.61)$$

This configuration is completely described by three parameters:  $\Delta\Re\lambda$ ,  $N$  and  $\Im\lambda$ . Reducing the number of parameters, we effectively simplify the system, however, the control over all signal parameters retains similarly. The integrals from the expressions for  $R(j, k)$ , eq. (3.92) reduce to:

$$\int_{\Gamma_j} \frac{\xi^k d\xi}{w(\xi)} = i \int_{-\Im\lambda}^{\Im\lambda} \frac{(\Re\lambda_j + it)^k dt}{w(\Re\lambda_j + it)}. \quad (6.62)$$

In the limit of  $\Im\lambda \rightarrow 0$  one can get the following:

$$\begin{aligned} \int_{\Gamma_j} \frac{\xi^k d\xi}{w(\xi)} &= i \int_{-\Im\lambda}^{\Im\lambda} \frac{(\Re\lambda_j + it)^k dt}{w_j(\Re\lambda_j + it) \sqrt{\Im\lambda_j^2 - t^2}} \\ &= 2i \int_0^{\pi/2} \frac{(\Re\lambda_j + i\Im\lambda \sin x)^k dx}{w_j(\Re\lambda_j + i\Im\lambda \sin x)} \xrightarrow{\Im\lambda \rightarrow 0} \pi i \frac{\lambda_j^k}{w_j(\lambda_j)}, \end{aligned} \quad (6.63)$$

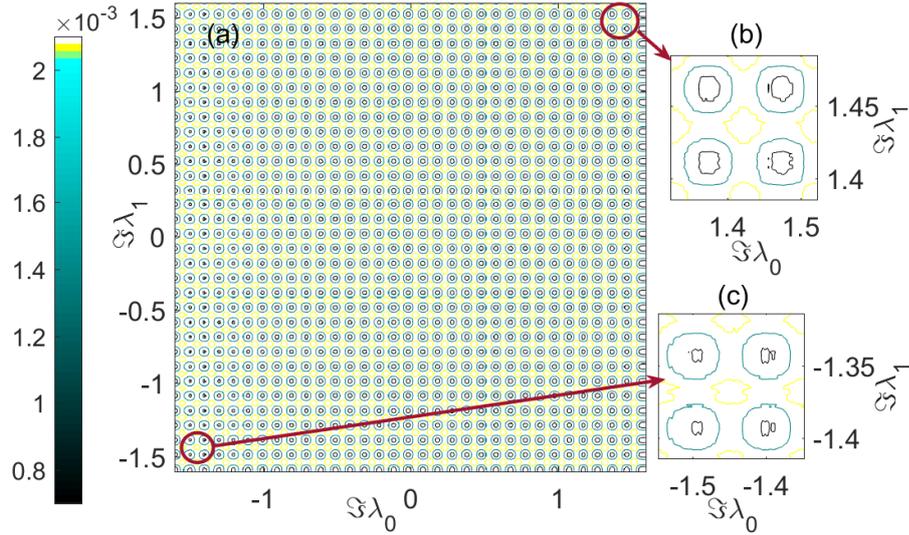


Figure 6.32: The simulation results of the transmission systems based on genus-1 periodic RHP solution: (a) 2D histogram of the received constellation of a 1024-QAM at  $Z = 680$  km and signal power  $P = -5$  dBm; with close up for two parts of the constellation attributed to (b) the highest and (c) the lowest signal power.

where

$$w_j(\lambda) = \prod_{\substack{i=-N/2 \\ i \neq j}}^{N/2} \sqrt{(\lambda - \lambda_i)(\lambda - \lambda_i^*)}. \quad (6.64)$$

The expression for  $R(j, k)$  can be therefore simplified to:

$$R(j, k) = \frac{\pi i j^{k+1} (\Delta \Re \lambda)^{k-N} (-1)^{N/2-j}}{(\mathcal{N}/2 - j)! (\mathcal{N}/2 + j)!}, \quad j = -N/2, \dots, -1, 1, \dots, N/2. \quad (6.65)$$

In zero-order approximation, the main frequency is  $C_1^f = -2\Delta \Re \lambda$ , and the signal period is given as  $T_{\text{norm}} = \pi / \Delta \Re \lambda$ . For a finite, nonzero signal power, that corresponds to a finite, nonzero value of  $\Im \lambda$ , the main spectrum points have to be (numerically) readjusted: using the equidistant grid as the initial guess, an optimisation problem is to be solved, where the main spectrum which renders the vector of  $C_j^f$ s closest to the desired values is to be reached.

The main spectrum structure chosen in our work allows us to manipulate the signal parameters, such that we can adjust this structure to get the desired values of power, period and bandwidth, leaving the freedom to change the associated phases. Three examples of a signal with three cuts with main spectrum  $\{-0.37 - 0.33i, 0.33i, 0.37 - 0.33i\}$  with different phase as the jump matrices are shown in fig. 6.33. As the signal waveforms look traditional for conventional optical communications, this approach appears to be more promising for hardware implementation.

The simulations are done for SMF link and ideal Raman amplification (i.e. the zero gain-loss profile). The presence of ASE is taken into account by adding the noise term to the r.h.s. of the NLS. The number of time samples per a single signal (containing 15 MPSK symbols), i.e. per one period, is 64. Each signal is then cyclically extended in time to the value of the dispersion-induced channel memory at  $Z = 1040$  km. These cyclic-prefix-appended signals are then cascaded to form a batch of several consecutive signals. To calculate each  $Q^2$ -factor and BER value,  $2^8$  of such batches were used.

An example of a signal carrying 15 MPSK symbols before adding a cyclic extension and the received signal after removing the cyclic extension is shown in fig. 6.34.

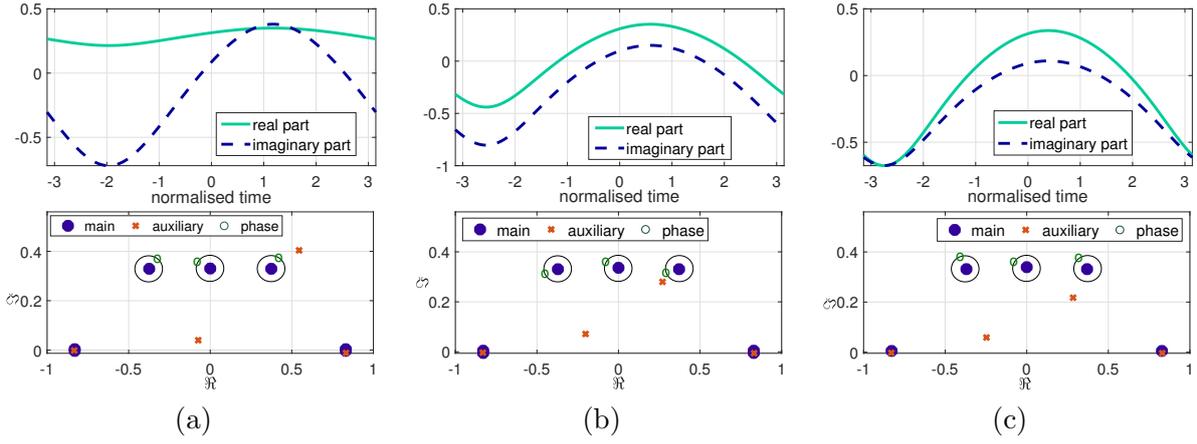


Figure 6.33: Examples of the generated signal with a fixed 3-cut main spectrum  $\lambda = \{-0.37 - 0.33i, 0.33i, 0.37 - 0.33i\}$  and difference associated phases: (a)  $\phi = \{\frac{3\pi}{8}, \frac{7\pi}{8}, \frac{3\pi}{8}\}$ , (b)  $\phi = \{\frac{9\pi}{8}, \frac{7\pi}{8}, \frac{9\pi}{8}\}$ , (c)  $\phi = \{\frac{5\pi}{8}, \frac{7\pi}{8}, \frac{5\pi}{8}\}$ . The top plots show the real and imaginary part of the signal, the bottom plots show the NF picture.

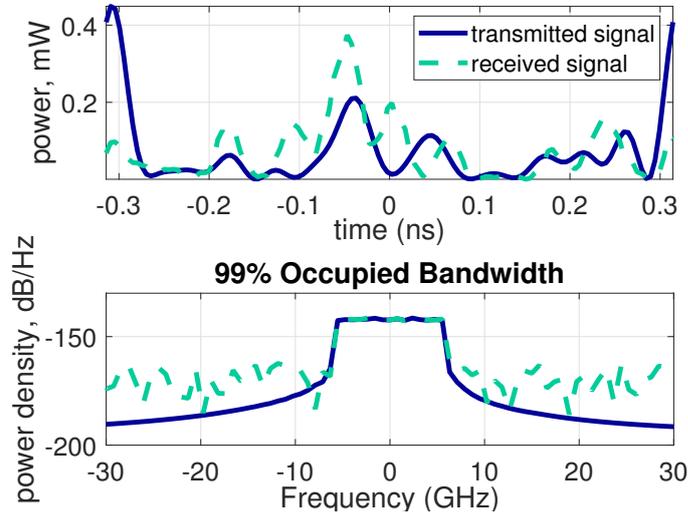


Figure 6.34: Example of the transmitted and received signal, generated from 15-cut RHP main spectrum, carrying 15 8PSK symbols over a link of length  $Z = 1080$  km.

The resulting  $Q^2$ -factor dependence against signal power is shown in fig. 6.35. The transmitted and received MPSK constellations are also shown in fig. 6.36 for two different signal powers. In fig. 6.36(a) plotted at the point of optimum signal power, an error-free communication is provided up to the accuracy of our simulations. We note that the achievable  $Q$ -factor degrades with the propagation distance, see fig. 6.37.

The average signal bandwidth is 9.8 GHz and the delivered data rate is 23.7 Gb/s up to 1040 km transmission, which shows a considerable improvement compared to main spectrum modulation from subsection 6.5.3.

Here we also note that the highest data rate for the soliton-based QAM transmission system reported was  $\sim 24$  Gb/s at 1000 km [46], such that our proof-of-concept system here has already reached that value, which actually reveals the potential of the PNFT-based approaches.

As we can observe in fig. 6.36, there is a deterministic rotation in the phase domain even after the back-propagation stage performed by using eq. (3.55). This phase rotation makes the received phases distribution asymmetric and can be remedied by means of a simple

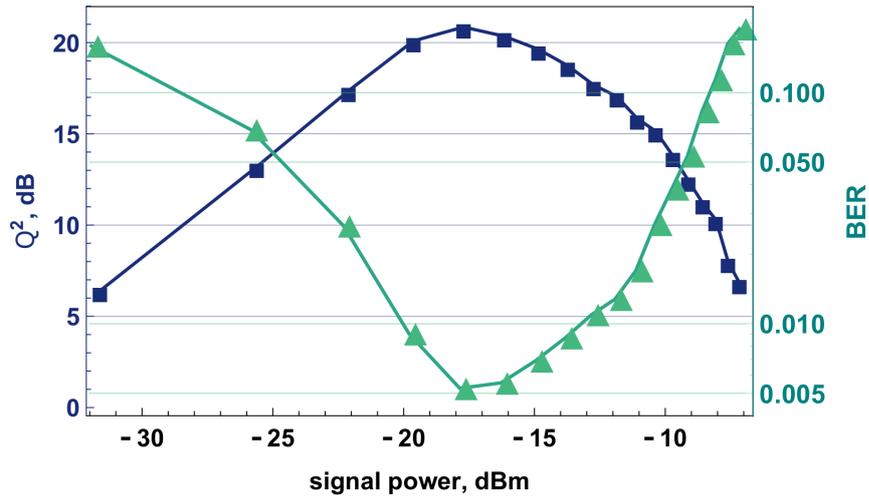


Figure 6.35: The  $Q^2_{EVM}$ -factor and BER as a function of signal power after  $Z = 1040$  km transmission for 15-cuts RHP-based transmission system. The optimum is achieved at  $P = -17$  dBm.

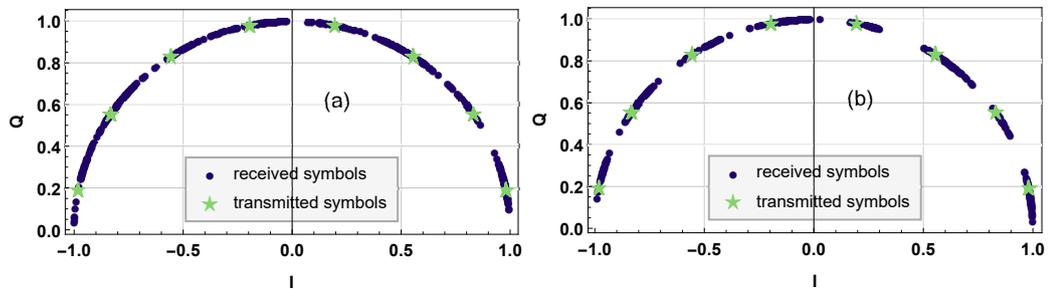


Figure 6.36: The received phase constellation after  $Z = 1000$  km transmission of an 8PSK auxiliary spectrum modulated signal at signal power (a)  $P = -6$  dBm, and (b)  $P = -17$  dBm.

blind or, as we have implemented in this work, a pilot-aided equaliser which improves the performance up to 1.5 dBm. The length of the preamble is inconsiderable compared to the signal length.

We note the SE of the proposed system is still small comparably to other state-of-the-art optical communication systems, basically due to the large time-bandwidth product value per baud in our RHP-generated time-domain signals. This issue leaves the space for the further improvements of the concept, studying the other main spectrum structures that would yield better values. However, as explained earlier, a more compact main NF spectrum where the points are chosen with the same real value and different imaginary parts leads to a signal with the same characteristics while carrying more data.

## 6.6 Communication systems discussion

This chapter is dedicated to various designs of NFT-based transmission systems. The conventional (vanishing) NFT provides more opportunities for data modulation, though is less convenient from the point of signal properties monitoring. Also, in the domain of conventional NFT, the available numerical processing methods are more mature. It is natural because this branch took more attention from researchers because of its closer connections to advances of linear communication.

In this thesis, the main focus is done for the advancing of existent conventional NFT approaches: first of all, by incorporation contour integrals as a replacement for multi-

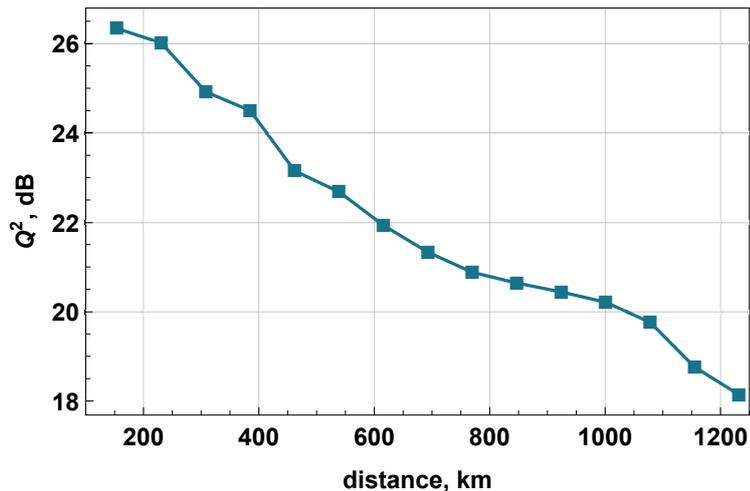


Figure 6.37: The  $Q^2_{\text{EVM}}$ -factor as a function of transmission distance at optimum signal power  $P = -17$  dBm.

eigenvalue communication. The noticeable part of the presented results is done within the recently introduced  $b$ -modulation approach, which provides control over the signal duration. For the first time, the full justification of the correspondence between bandlimited scattering function  $b$  and finitely supported optical signal  $q$  is proven involving the powerful and flexible tool of the RHP. Additionally, it allows establishing a way to identify the eigenvalues which maintain the signal localisation. Several publications are dedicated to advancing  $b$ -modulated signal with discrete spectrum modulation.

The development of periodic NFT-based transmission is a rather challenging task, because there are fewer similarities both with linear systems, and conventional NFT, so the achievements of neighbouring areas cannot be directly applied here. However, the results presented here have, first, fundamental value, since the finite-band theory for focusing NLS is less developed, despite its clear application importance. Additionally, the transition from arbitrary finite-band solution to exactly periodic one with a study of signal properties relations to the main spectrum structure makes these systems more practical of design of predictable waveforms.

However, we have to admit that the data rates of the existent systems must be improved significantly to meet the growing data traffic demand, and to showcase the relevance of NFT to the efficient mitigation of nonlinear impairments in optical communications.

## Chapter 7

# Overall discussion and conclusion

The NFT-based communication concept provides two opportunities to model the channel and apply the linearising mapping: vanishing or periodic boundary conditions, see chapter 3. Despite that the general problem of the constructing periodic solution of NLS is not solved, for both cases, it is possible to build one-to-one correspondence between the NF image and the time-domain waveform. These methods provide different flexibility in controlling signal parameters and degrees of freedom for data modulation. So it is generally left on a transmission system designer how the use them most effectively. The general observation is that periodic NFT allows fewer possibilities for the information modulation, because the NF spectrum objects are discrete and their number dramatically influence the complexity of the system. In turn, the conventional (vanishing) NFT suffers from the inconvenient waveform and expanding processing window.

It seems that the application of  $b$ -modulation, studied in section 6.4, but for two polarisation case, allowed to reach the record-breaking 400 Gbps transmission on 1000 km [146, 147] is close to the maximum achievable data rate for the NFT-based transmission at least in its conventional (vanishing) format (it may be further enriched by discrete modes modulation, as shown in this thesis, however, it is clear that with no significant gain in data rate). The perspectives of the periodic approach are less studied up to date. The Riemann-Hilbert problem method, presented in this thesis in section 3.2, looks like a powerful and universal tool for these purposes. The idea of this approach is principally different from what was studied before (contours in the complex planes, jump matrices, etc.), which makes the effective study rather challenging. However, the transmission systems which we managed to construct, provide rather miserable values of data rate (maximum in 24 Gbps at 1000 km) and SE [SJ1]. These low values cannot be sufficiently convincing in attracting more research resources for this topic.

This thesis is dedicated to furthering and developing the concept of NFT applied to long-distance optical communications. In this work, I, for the first time, demonstrated the performance of numerical methods for discrete eigenvalues evaluation (contour integrals approach in chapter 5), basing on the various testing examples, containing several eigenvalues, closely located eigenvalues and even degenerate eigenvalues; developed the communication systems employing degrees of freedom which are available in the finite-band NLS solutions; the performance and signal characteristics for such systems are studied in details.

The overall research in this area reveals different severe difficulties in the effective implementing of the NFT ideology to the transmission. I would like to discuss them here. The integrability of NLS because of the lucky coincidence of competing terms, is actually ruined by the practical design of the communication systems. Apart from the optical fibre, it consists of transmitted and receivers, optical-to-electrical converters (and back), analogue-to-digital converters (and back), amplifiers and repeaters. All these parts

introduce their imperfections and noise.

In the linear approach, all linear deviations of Gaussian stochasticity are likely to lead to the similarly linear deviations and Gaussian stochasticity, that facilitate the development of the compensation techniques. So in the linear approach, all deviation from the pure NLS lead to manageable effects. In NFT approach, the deviations produce dramatic consequences for the very idea of integrability. Frankly speaking, we cannot easily ignore this deviation, unless the deviations are extremely small (what is never the case).

Traditionally, the compensation techniques for transceivers nonlinearity are developed under the assumption of the linear channel. Therefore, the achievements of conventional telecommunication are not rather applicable for NFT-based transmission, which put us well back in the competition for higher data rate. One of these deviations is an amplification in the fibre accompanied by the amplifier noise. Both of them lead to a highly noisy picture of the constellations in NF domain. One of the parts of my research, leading by Dr Maryna Pankratova [J5, CC1, CC7], effectively confirmed the intuitively anticipated fact that the noise caused in NF domain by feeding the time domain with ASE noise, has complicated, nonlinear and input-dependent nature, which makes the usage of such systems rather impractical.

Another challenging part in the developing of the numerical methods for both processing on the Tx and Rx. The first problem is that the available methods, studied in details in chapter 4, are still rather slow and used only for off-line processing, despite the much efforts done in this direction. Another problem is accuracy dependency on signal power, coupled with the sampling requirements. It is widely reported that NFT processing require signal oversampling to ensure acceptable accuracy. Finally, the field still requires reliable, fast and simple methods for NFT processing

From the hardware side, NFT-based transmission involves untypical waveforms and structures, requiring more expensive and less efficient components. Together with that, the effective online implementation of the NFT-based processing requires the development of specially designed components (possibly, partly or all-optical), what can be the topic of yet another research project, due to many degrees of freedom and decision points on the trade-off between accuracy, computational efforts and final cost.

Answering the question 'Why NFT does not work in optcomm?' I usually refer to the lack of engineering expertise among people who work in the area, and difficulties in the overcoming of mathematical and technical challenges. The complicated mathematics makes this area too abstract and scary for the engineering community. Finally, it leads to the current situation, when people involved in this research are mainly with theoretical/mathematical/signal processing background, and there is a noticeable demand in experimental telecommunication excellence. Another point of view and principally different research approach could be salvatory for the topic.

Moreover, I identify the tendency to abandoning of this topic by main research groups. Despite that NFT-oriented sessions were introduces at ECOC 2018 and 2019 (what means a noticeable level of appreciation and acceptance in optcomm community), people start reorienting from this topic to something else. Majority of groups are focused now on the numerical methods, rather on the developing of the transmission systems. The general situation is not quite optimistic. As I understand, something similar happened with the soliton-based transmission decades ago.

Here I also want to highlight the perspective of the listed results apart from the applications for optical communications. It includes a large amount of purely mathematical or numerical findings. First, a numerical solution of the ZSP applies to other areas, like coupled-mode theory, or design of fibre Bragg grating [38] - important on-line device of dispersion compensation in the fibre. There are also applications of ZSP for *LCGR* transmission lines. The idea of contour integrals for the location of the eigenvalues can be

generalised to any other isolated objects in the complex plane of the spectral parameters (e.g., for the PNFT main spectrum identification), or even beyond the IST concept. There are some adaptations of the technique could be required, however, I anticipate similar advantages comparably to other zero-search methods.

The findings within the RHP approach are generalisable in various dimensions. The path of the constructing periodic solution starting from the finite-band one is important from the fundamental perspective, as the general problem of construction of arbitrary periodic solution of NLS is not solved, especially the case of focusing choice of signs. In general, the focusing NLS is less studied from the point of finite-band IST because of more complex structures in NF domain. Moreover, as a temporal and distance parameters  $\tau$  and  $\zeta$  enter the RHP solutions, it is easy to derive the solution for any value, particularly, for large- $\zeta$  asymptotes, what is valuable for such applications, as deep and shallow water waves dynamics. Finally, the part of the determination of eigenvalues which maintain the localisation of  $b$ -modulated signal, alongside with completing the one-to-one correspondence between properties of  $b$ -function and NLS solution, is a remarkable showcase of the flexibility of the RHP method for proving practical properties of the studied NLS solutions.

In general, this research area effectively provided for me the quality, high-level training in studying of the interdisciplinary subject. Ordering from more fundamental to more practical fields, I would start with mathematical physics, especially nonlinear and coherent structures science. I found it being a bit of a miracle skill, how researchers work with the equations and terms within them, to suppress or give birth to specific solutions, their asymptotes or properties. My work also involved plenty of algebraic, asymptotic and perturbative analysis of the ODE and PDE solutions. I would say that the practising in analytical derivations is an effective brain training tool, and I found doing it both enjoyable and refreshing. I also involved some knowledge from communication and information theory science. Due to my theoretical physics background, it was rather complicated to communicate and interact properly with engineering people, books and journals.

Also, the noteworthy part of my doctoral training is a research project planning, project- and time-managements. I took the in-hand experience in paper publication, with all review, correcting, resubmission processes. The fact that my first publication was ready in the first half-year of my three-year course, allowed me to practice from the other side of the publishing process, I mean being a research paper reviewer. The experience in attending, presenting and networking on the large topical meetings, like ECOC and CLEO-Europe, also contributed to my professional development and confidence. All abovementioned allow me to conclude that my training as a researcher is successful and complete.

# Bibliography

- [1] M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur. The inverse scattering transform-Fourier analysis for nonlinear problems. *Studies in Appl. Math.*, 53(4):249–315, 1974.
- [2] M. J. Ablowitz and H. Segur. *Solitons and the Inverse Scattering Transform*. SIAM, 1981.
- [3] C. Affane-Aji, N. Agarwal, and N. K. Govil. Location of zeros of polynomials. *Mathematical and Computer Modelling*, 50(1-2):306–313, 2009.
- [4] G. P. Agrawal. *Nonlinear Fiber Optics*. Academic Press, 2013.
- [5] N. Akhmediev, A. Ankiewicz, and M. Taki. Waves that appear from nowhere and disappear without a trace. *Phys. Lett. A*, 373:675–678, 2009.
- [6] T. Aktosun. Inverse scattering transform and the theory of solitons. *Encyclopedia of Complexity and Systems Science*, pages 4960–4971, 2009.
- [7] A. Amari, X. Lin, O. A. Dobre, R. Venkatesan, and A. Alvarado. A machine learning-based detection technique for optical fiber nonlinearity mitigation. *IEEE Photonics Technology Letters*, 31(8):627–630, 2019.
- [8] J. D. Ania-Castañón, T. J. Ellingham, R. Ibbotson, X. Chen, L. Zhang, and S. K. Turitsyn. Ultralong Raman fiber lasers as virtually lossless optical media. *Physical review letters*, 96(2):023902, 2006.
- [9] J. D. Ania-Castañón, V. Karalekas, P. Harper, and S. K. Turitsyn. Simultaneous spatial and spectral transparency in ultralong fiber lasers. *Physical review letters*, 101(12):123903, 2008.
- [10] V. Aref. Control and detection of discrete spectral amplitudes in nonlinear Fourier spectrum. *preprint*, 2016. arXiv:1605.06328v1.
- [11] V. Aref, S. T. Le, and H. Buelow. Modulation over nonlinear Fourier spectrum: Continuous and discrete spectrum. *Journal Lightwave Technologies*, 36:1289, 2018.
- [12] V. Aref, S. T. Le, and H. Buelow. An efficient Nonlinear Fourier Transform algorithm for detection of eigenvalues from continuous spectrum. In *Optical Fiber Communication Conference (OFC) 2019*, page M1L.5. Optical Society of America, 2019.
- [13] A. Barone, F. Esposito, C. J. Magee, and A. C. Scott. Theory and applications of the sine-Gordon equation. *La Rivista del Nuovo Cimento (1971-1977)*, 1(2):227–267, 1971.

- [14] E. D. Belokolos, A. I. Bobenko, V. Z. Enolski, A. R. Its, and V. B. Matveev. *Algebro-geometric Approach in the Theory of Integrable Equations*. In *Springer Series in Nonlinear Dynamics*. Springer, 1994.
- [15] A. Bjarklev. *Optical fiber amplifiers: design and system applications*. Artech House Boston-London, 1993.
- [16] R. P. Boas. *Entire Functions*. Academic Press Inc., 1954.
- [17] A. I. Bobenko and L. A. Bordag. Periodic multiphase solutions of the Kadomsev-Petviashvili equation. *J. of Physics A: Mathematical and General*, 22:1259–1274, 1989.
- [18] G. Boffetta and A. Osborne. Computation of the direct scattering transform for the nonlinear Schroedinger equation. *J. Comput. Phys.*, 102:252–264, 1992.
- [19] S. Burtsev, R. Camassa, and I. Timofeyev. Numerical algorithms for the direct spectral transform with application to nonlinear schrodinger type systems. *Journal of Computational Physics*, 147:166–186, 1998.
- [20] F. Caballero, D. Ives, C. Laperle, D. Charlton, Q. Zhuge, M. O’Sullivan, and S. Savory. Machine learning based linear and nonlinear noise estimation. *Journal of Optical Communications and Networking*, 10(10):D42–D51, 2018.
- [21] A. Calini and C. M. Schober. Characterizing JONSWAP rogue waves and their statistics via inverse spectral data. *Wave Motion*, 71:5–17, 2017.
- [22] J. C. Cartledge, F. P. Guiomar, F. R. Kschischang, G. Liga, and M. P. Yankov. Digital signal processing for fiber nonlinearities. *Opt. Express*, 25(3):1916–1936, 2017.
- [23] S. Chimmalgi, P. J. Prins, and S. Wahls. Fast nonlinear Fourier transform algorithms using higher order exponential integrators. *IEEE Access*, 7:145161–145176, 2019.
- [24] A. Chowdhury and R. J. Essiambre. Optical phase conjugation and pseudolinear transmission. *Optics letters*, 29(10):1105–1107, 2004.
- [25] S. Civelli, E. Forestieri, and M. Secondini. Impact of discretization and boundary conditions in nonlinear frequency-division multiplexing. In *18th Italian National Conference on Photonic Technologies (Fotonica 2016)*, 2016.
- [26] S. Civelli, E. Forestieri, and M. Secondini. Why noise and dispersion may seriously hamper nonlinear frequency-division multiplexing. *IEEE Photon. Tech. Lett.*, 29(16):1332–1335, 2017.
- [27] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.*, 19:297–301, 1965.
- [28] R. Dar and P. J. Winzer. On the limits of digital back-propagation in fully loaded WDM systems. *IEEE Photonics Technology Letters*, 28(11):1253–1256, 2016.
- [29] B. Deconinck, M. Heil, A. Bobenko, M. van Hoeij, and M. Schmies. Computing Riemann theta functions. *Math. Comp.*, 73:1417–1442, 2004.
- [30] L. M. Delves and J. N. Lyness. A numerical method for locating the zeros of an analytic function. *Math. Comp.*, 21:543–560, 1967.

- [31] S. A. Derevyanko and J. E. Prilepsky. Random input problem for the nonlinear schrödinger equation. *Physical Review E*, 78(4):046610, 2008.
- [32] S. A. Derevyanko, J. E. Prilepsky, and S. K. Turitsyn. Capacity estimates for optical transmission based on the nonlinear Fourier transform. *Nature Commun*, 7(12710):307, 2016.
- [33] E. Desurvire, D. Bayart, B. Desthieux, and S. Bigo. *Erbium-doped fiber amplifiers: Device and System Developments*, volume 2. Wiley-Interscience New York, NY, 2002.
- [34] C. Dorrer, C. R. Doerr, I. Kang, R. Ryf, J. Leuthold, and P. J. Winzer. Measurement of eye diagrams and constellation diagrams of optical sources using linear optics and waveguide technology. *Journal of Lightwave Technology*, 23(1):178, 2005.
- [35] A. D. Ellis, M. Tan, M. A. Iqbal, M. A. Z. Al-Khateeb, V. Gordienko, et al. 4 Tb/s transmission reach enhancement using  $10 \times 400$  Gb/s super-channels and polarization insensitive dual band optical phase conjugation. *Journal of lightwave technology*, 34(8):1717–1723, 2016.
- [36] R. J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel. Capacity limits of optical fiber networks. *J. Lightwave Technol*, 28:662–701, 2010.
- [37] L. D. Faddeev and L. A. Takhtajan. *Hamiltonian Methods in the Theory of Solitons*. Springer Series in Soviet Mathematics. Springer, 1987.
- [38] L. L. Frumin, O. V. Belai, E. V. Podivilov, and D. A. Shapiro. Efficient numerical method for solving the direct Zakharov-Shabat scattering problem. *J. Opt. Soc. Am. B*, 32:290–295, 2015.
- [39] S. Gaiarin, F. Da Ros, N. De Renzis, E. P. da Silva, and D. Zibar. Dual-polarization NFDM transmission using distributed Raman amplification and NFT-domain equalization. *IEEE Photonics Technology Letters*, 30(22):1983–1986, 2018.
- [40] S. Gaiarin, A. M. Perego, E. P. da Silva, F. Da Ros, , and D. Zibar. Dual-polarization nonlinear Fourier transform-based optical communication system. *Optica*, 5(3):263–270, 2018.
- [41] F. J. García-Gómez. Communication using eigenvalues of higher multiplicity of the nonlinear Fourier transform. *Journal of Lightwave Technology*, 36(23):5442–5450, 2018.
- [42] F. J. García-Gómez and V. Aref. Statistics of the nonlinear discrete spectrum of a noisy pulse. *J. Lightwave Technol.*, 37(14):3563–3570, 2019.
- [43] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R.M. Miura. Method for solving the Korteweg-de Vries equation. *Physical Review Letters*, 19:1095–1097, 1967.
- [44] G. Gasper and M. Rahman. *Basic Hypergeometric Series*. Cambridge University Press, 2nd edition, 2004.
- [45] E. Giacomidis, A. Matin, J. Wei, N. J. Doran, L. P. Barry, and X. Wang. Blind nonlinearity equalization by machine-learning-based clustering for single-and multi-channel coherent optical OFDM. *Journal of Lightwave Technology*, 36(3):721–727, 2018.

- [46] T. Gui, C. Lu, A. P. T. Lau, and P. K. A. Wai. High-order modulation on a single discrete eigenvalue for optical communications based on nonlinear Fourier transform. *Optics express*, 25(17):20286–20297, 2017.
- [47] T. Gui, G. Zhou, C. Lu, A. P. T. Lau, and S. Wahls. Nonlinear frequency division multiplexing with b-modulation: Shifting the energy barrier. *Optics Express*, 26:27978–27990, 2018.
- [48] F. P. Guiomar, J. D. Reis, A. L. Teixeira, and A. N. Pinto. Digital postcompensation using Volterra series transfer function. *IEEE Photonics Technology Letters*, 23(19):1412–1414, 2011.
- [49] B. Guo, L. Ling, and Q. P. Liu. Nonlinear Schrödinger equation: Generalized Darboux transformation and rogue wave solutions. *Phys. Rev. E*, 85:026607, 2012.
- [50] S. Hari, F. Kschischang, and M. Yousefi. Multi-eigenvalue communication via the nonlinear Fourier transform. In *27th Biennial Symposium on Communications (QBSC)*, pages 92–95, 2014.
- [51] S. Hari and F. R. Kschischang. Bi-directional algorithm for computing discrete spectral amplitudes in the NFT. *Journal of Lightwave Technology*, 34(15):3529–3537, 2016.
- [52] S. Hari, M. I. Yousefi, and F. R. Kschischang. Multieigenvalue communication. *J. Lightwave Technol.*, 34:3110–3117, 2016.
- [53] A. Hasegawa and Y. Kodama. *Solitons in optical communications*. Oxford University Press, USA, 1995.
- [54] A. Hasegawa and T. Nyu. Eigenvalue communication. *J. Lightwave Technol.*, 11(3):395–399, 1993.
- [55] E. Iannone, F. Matera, A. Mecozzi, and M. Settembre. *Nonlinear optical communication networks*. Wiley, 1998.
- [56] E. Ip and J. Kahn. Compensation of dispersion and nonlinear impairments using digital backpropagation. *J. Lightwave Technol.*, 26:3416–3425, 2008.
- [57] M. A. Iqbal, M. A. Z. Al-Khateeb, L. Krzczanowicz, I. D. Phillips, P. Harper, and W. Forsyiaik. Linear and nonlinear noise characterisation of dual stage broadband discrete Raman amplifiers. *Journal of Lightwave Technology*, 37(14):3679–3688, 2019.
- [58] A. R. Its and V. B. Matveev. The periodic Korteweg-deVries equation. *Funct. Anal. Appl.*, 9, 1975.
- [59] O. R. Its and V. P. Kotlyarov. Explicit formulas for the solutions of a nonlinear Schrodinger equation. *Doklady Akad. Nauk Ukrainian SSR, ser. A*, 10:965–968, 1976.
- [60] M. K. Jain. *Numerical Methods For Scientific And Engineering Computation*. New Age International, 2003.
- [61] S. L. Jansen, D. Van den Borne, B. Spinnler, S. Calabro, H. Suche, P. M. Krummrich, W. Sohler, G.-D. Khoe, and H. De Waardt. Optical phase conjugation for ultra long-haul phase-shift-keyed transmission. *Journal of Lightwave Technology*, 24(1):54, 2006.

- [62] L. W. Johnson and D. R. Scholz. On Steffensen's method. *SIAM Journal on Numerical Analysis*, 5(2):296–302, 1968.
- [63] M. Kamalian, J. E. Prilepsky, S. T. Le, and S. K. Turitsyn. Periodic nonlinear Fourier transform for fiber-optic communications, parts i-ii. *Opt. Express*, 24:18353–18369, 2016.
- [64] M. Kamalian, J. E. Prilepsky, S. T. Le, and S. K. Turitsyn. On the design of NFT-based communication systems with lumped amplification. *Journal of Lightwave Technology*, 35(24):5464–5472, 2017.
- [65] A. M. Kamchatnov. New approach to periodic solutions of integrable equations and nonlinear theory of modulational instability. *Physics Reports*, 286(4):199–270, 1997.
- [66] R. Kashyap. *Fiber Bragg Gratings*. Academic press, 2009.
- [67] D. J. Kedziora, A. Ankiewicz, and N. Akhmediev. Second-order nonlinear schrödinger equation breather solutions in the degenerate and rogue wave limits. *Phys. Rev. E*, 85, 2012.
- [68] F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau. An optical communication's perspective on machine learning and its applications. *Journal of Lightwave Technology*, 37(2):493–516, 2019.
- [69] F. N. Khan, C. Lu, and A. P. T. Lau. Machine learning methods for optical communication systems. In *Signal Processing in Photonic Communications*, page SpW2F3. Optical Society of America, 2017.
- [70] N. Kikuchi, K. Sekine, and S. Sasaki. Multilevel signalling for high-speed optical transmission. In *2006 European Conference on Optical Communications*, 2006.
- [71] D. J. Korteweg and G. de Vries. On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 39(240):422–443, 1895.
- [72] V. Kotlyarov and D. Shepelsky. Planar unimodular Baker-Akhiezer function for the nonlinear Schrödinger equation. *Ann. Math. Sci. Appl.*, 2:343–384, 2017.
- [73] V. P. Kotlyarov. Periodic problem for the nonlinear Schrödinger equation. *Problems Math. Phys. Funct. Anal.*, 1:121–131, 1976.
- [74] D. Kraus, G. Biondini, and G. Kovačič. The focusing Manakov system with nonzero boundary conditions. *Nonlinearity*, 28(9):3101, 2015.
- [75] P. Kravanja, T. Sakurai, and M. van Barel. On location clusters of zeros of analytic functions. *BIT*, 39(4):646–682, 1999.
- [76] I. M. Krichever. The spectral theory of two-dimensional periodic operators and applications. *Russ. Math. Surv.*, 44:121–184, 1989.
- [77] L. Krzaczanowicz, M. A. Iqbal, I. D. Phillips, M. Tan, P. Skvortcov, P. Harper, and W. Forsyiaik. Low transmission penalty dual-stage broadband discrete Raman amplifier. *Opt. Express*, 26(6):7091–7097, 2018.
- [78] G. Lamb. *Elements of soliton theory*. Wiley-Interscience, New York, 1980.

- [79] P. D. Lax. Integrals of nonlinear equations of evolution and solitary waves. *Communications on pure and applied mathematics*, 21(5):467–490, 1968.
- [80] S. Le, V. Aref, and H. Buelow. Nonlinear signal multiplexing for communication beyond the Kerr nonlinearity limit. *Nature Photonics*, 11:570, 2017.
- [81] S. T. Le, V. Aref, and H. Buelow. High speed precompensated nonlinear frequency-division multiplexed transmissions. *J. Lightwave Technol.*, 36(6):1296–1303, 2018.
- [82] S. T. Le and H. Buelow.  $64 \times 0.5$  Gbaud nonlinear frequency division multiplexed transmissions with high order modulation formats. *J. Lightwave Technol.*, 35(17):3692–3698, 2017.
- [83] S. T. Le, I. Philips, J. Prilepsky, P. Harper, N. Doran, A. D. Ellis, and S. Turitsyn. First experimental demonstration of nonlinear inverse synthesis transmission over transoceanic distances. In *in Optical Fiber Communication Conference (OFC), OSA Technical Digest, Anaheim, USA*, 2016.
- [84] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn. Nonlinear inverse synthesis for high spectral efficiency transmission in optical fibers. *Optics Express*, 22:26270, 2014.
- [85] S. T. Le, J. E. Prilepsky, and S. K. Turitsyn. Nonlinear inverse synthesis technique for optical links with lumped amplification. *Optics express*, 23(7):8317–8328, 2015.
- [86] S. T. Le, K. Schuh, F. Buchali, and H. Bülow. 100 Gbps b-modulated nonlinear frequency division multiplexed transmission. In *Proc. Optical Fiber Communication Conference (OFC)*, page W1G.6., 2018.
- [87] B. Ya. Levin. *Lectures on Entire Functions*, volume 150 of *Transl. Mathem. Monographs*. AMS, 1996.
- [88] X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, and G. Li. Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing. *Optics Express*, 16(2):880–888, 2008.
- [89] I. Lima, V. Grigoryan, M. O’Sullivan, and C. Menyuk. Computational complexity of nonlinear transforms applied to optical communications systems with normal dispersion fibers. In *Proc. IEEE Photonics Conference*, pages 277–278, 2015.
- [90] Ivan T Lima, Thiago DS DeMenezes, Vladimir S Grigoryan, Maurice O’Sullivan, and Curtis R Menyuk. Nonlinear compensation in optical communications systems with normal dispersion fibers using the nonlinear fourier transform. *Journal of Lightwave Technology*, 35(23):5056–5068, 2017.
- [91] L. Liu, L. Li, Y. Huang, K. Cui, Q. Xiong, F. N. Hauske, C. Xie, and Y. Cai. Intrachannel nonlinearity compensation by inverse Volterra series transfer function. *Journal of Lightwave Technology*, 30(3):310–316, 2011.
- [92] A. Lord, A. Soppera, and A. Jacquet. The impact of capacity growth in national telecommunications networks. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2062):20140431, 2016.
- [93] J. N. Lyness. Differentiation formulas for analytic functions. *Mathematics of Computation*, 22(102):352–362, 1968.
- [94] J. N. Lyness and L. M. Delves. On numerical contour integration round a closed contour. *Mathematics of computation*, 21(100):561–577, 1967.

- 
- [95] J. N. Lyness and C. B. Moler. Numerical differentiation of analytic functions. *SIAM Journal on Numerical Analysis*, 4(2):202–210, 1967.
- [96] Y. Ma and M. J. Ablowitz. The periodic cubic Schrödinger equation. *Stud. Appl. Math.*, 65:113–158, 1981.
- [97] Y.-C. Ma. *Studies of the cubic Schrödinger equation*. PhD thesis, Princeton Univ., 1977.
- [98] S. V. Manakov. On the theory of two-dimensional stationary self-focusing of electromagnetic waves. *Soviet Physics-JETP*, 38(2):248–253, 1974.
- [99] V. B. Matveev. 30 years of finite-gap integration theory. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 366:837–875, 2008.
- [100] K. T. R. McLaughlin and P. V. Nabelek. A Riemann-Hilbert problem approach to periodic infinite gap Hill operators and the Korteweg-de Vries equation. *arXiv preprint arXiv:1810.07818*, 2018.
- [101] S. Medvedev, I. Chekhovskoy, I. Vaseva, and M. Fedoruk. Fast computation of the direct scattering transform by fourth order conservative multi-exponential scheme. *arXiv preprint arXiv:1909.13228*, 2019.
- [102] S. Medvedev, I. Vaseva, I. Chekhovskoy, and M. Fedoruk. Numerical algorithm with fourth-order accuracy for the direct Zakharov-Shabat problem. *Optics Letters*, 44(9):2264–2267, 2019.
- [103] S. B. Medvedev, I. A. Vaseva, I. S. Chekhovskoy, and M. P. Fedoruk. A novel fourth-order difference scheme for the direct Zakharov-Shabat problem. In *The European Conference on Lasers and Electro-Optics*, page ci.1.5, 2019.
- [104] D. E. Muller. A method for solving algebraic equations using an automatic computer. *Mathematical Tables and Other Aids to Computation*, 10(56):208–215, 1956.
- [105] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore. An overview on application of machine learning techniques in optical networks. *IEEE Communications Surveys Tutorials*, 21(2):1383–1408, 2019.
- [106] A. Napoli, Z. Maalej, V. A. J. M. Sleiffer, M. Kuschnerov, D. Rafique, E. Timmers, B. Spinnler, T. Rahman, L. D. Coelho, and N. Hanik. Reduced complexity digital back-propagation methods for optical communication systems. *Journal of lightwave technology*, 32(7):1351–1362, 2014.
- [107] S. Olver. RHPackage for MATHEMATICA. <https://github.com/dlfivefifty/RHPackage>.
- [108] S. Olver. Riemannhilbert.jl. <https://github.com/JuliaHolomorphic/RiemannHilbert.jl>.
- [109] S. Olver. A general framework for solving Riemann-Hilbert problems numerically. *Numerische Mathematik*, 122:305–340, 2012.
- [110] A. R. Osborne. *Nonlinear Ocean Waves and the Inverse Scattering Transform*. Academic Press, 2010.
- [111] J. Pan and C.-H. Cheng. Nonlinear electrical compensation for the coherent optical OFDM system. *Journal of Lightwave Technology*, 29(2):215–221, 2010.

- [112] V. E. S. Parahyba, J. D. Reis, S. M. Ranzini, E. O. Schneider, E. S. Rosa, et al. Performance against implementation of digital backpropagation for high-speed coherent optical systems. *Electronics Letters*, 51(14):1094–1096, 2015.
- [113] K. V. Peddanarappagari and M. Brandt-Pearce. Volterra series transfer function of single-mode fibers. *Journal of lightwave technology*, 15(12):2232–2241, 1997.
- [114] I. D. Phillips, M. Tan, M. F. C. Stephens, M. E. McCarthy, E. Giacomidis, S. Sygletos, P. Rosa, S. Fabbri, S. T. Le, T. Kanesan, et al. Exceeding the nonlinear-Shannon limit using Raman laser based amplification and optical phase conjugation. In *2014 Optical Fiber Communications Conference (OFC)*, 2014.
- [115] J. C. Portinari. An inverse scattering transform for potentials of compact support. *J. Math. Phys.*, 19(10):2100–2102, 1978.
- [116] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. LU decomposition and its applications. *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, pages 34–42, 1992.
- [117] J. E. Prilepsy, S. A. Derevyanko, K. J. Blow, I. Gabitov, and S. K. Turitsyn. Nonlinear inverse synthesis and eigenvalue division multiplexing in optical fiber channels. *Phys. Rev. Lett.*, 113, 2014.
- [118] B. Prinari, M. J. Ablowitz, and G. Biondini. Inverse scattering transform for the vector nonlinear Schrödinger equation with nonvanishing boundary conditions. *Journal of mathematical physics*, 47(6):063508, 2006.
- [119] B. Prinari, G. Biondini, and A. D. Trubatch. Inverse scattering transform for the multi-component nonlinear Schrödinger equation with nonzero boundary conditions. *Studies in Applied Mathematics*, 126(3):245–302, 2011.
- [120] D. Rafique, J. Zhao, and A. D. Ellis. Digital back-propagation for spectrally efficient WDM 112 Gbit/s PM m-ary QAM transmission. *Optics Express*, 19(6):5219–5224, 2011.
- [121] S. M. Ranzini, V. E. Parahyba, T. Vilela, E. O. Schneider, J. B. Tardelli, J. C. R. F. Oliveira, E. P. Lopes, T. C. Lima, J. D. Reis, J. R. F. Oliveira, and G. Fraidenraich. Digital back-propagation ASIC design for high-speed coherent optical system. In *2015 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC)*, 2015.
- [122] J. Renaudier, A. Arnould, A. Ghazisaeidi, D. Le Gac, P. Brindel, E. Awwad, M. Makhsyan, K. Mekhazni, F. Blache, B. Duval, and D. Lanteri. Recent advances in 100+ nm ultra-wideband fiber-optic transmission systems using semiconductor optical amplifiers. *Journal of Lightwave Technology*, 2020.
- [123] C. Rogers, C. Rogers, and W. K. Schief. *Bäcklund and Darboux transformations: geometry and modern applications in soliton theory*, volume 30. Cambridge University Press, 2002.
- [124] J. S. Russell. Report on waves. *Report of the 14th meeting of the British Association for the Advancement of Science*, page 311–390, 1845.
- [125] P. Sacks. An inverse problem in coupled mode theory. *J. Math. Phys.*, pages 1699–1710, 2004.

- [126] J. Satsuma and N. Yajima. Initial value problems of one-dimensional self-modulation of nonlinear waves in dispersive media. *Progr. Theoretical Phys. Suppl.*, 55(1):284–306, 1974.
- [127] C. E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, 1949.
- [128] T. S. R. Shen and A. P. T. Lau. Fiber nonlinearity compensation using extreme learning machine for DSP-based coherent communication systems. In *16th Opto-Electronics and Communications Conference*, pages 816–817, 2011.
- [129] A. Sidi. Generalization of the secant method for nonlinear equations. *Applied Mathematics E-Notes*, 8:115–123, 2008.
- [130] A. O. Smirnov. Periodic two-phase rogue waves. *Math. Notes*, 94:897–907, 2013.
- [131] E. R. Tracy and H. H. Chen. Nonlinear self-modulation: an exactly solvable model. *Phys. Rev. A*, 37:815–839, 1988.
- [132] L. N. Trefethen and D. Bau. *Numerical linear algebra*, volume 50. Siam, 1997.
- [133] T. Trogdon and S. Olver. Numerical inverse scattering for the focusing and defocusing nonlinear Schrödinger equations. *Proc R Soc A*, 469:284–306, 2013.
- [134] T. Trogdon and S. Olver. *Riemann-Hilbert Problems, Their Numerical Solution, and the Computation of Nonlinear Special Functions*. SIAM, 2016.
- [135] S. K. Turitsyn, J. E. Prilepsky, S. T. Le, S. Wahls, L. L. Frumin, M. Kamalian, and S. A. Derevianko. Nonlinear Fourier transform for optical data processing and transmission: Advances and perspectives. *Optica*, 3(4):307–322, 2017.
- [136] V. Vaibhav. Nonlinear Fourier transform of time-limited and one-sided signals. *J. Phys. A: Math. Theor.*, 51(42):425201, 2018.
- [137] V. Vaibhav and S. Wahls. Introducing the fast inverse NFT. In *2017 Optical Fiber Communications Conference and Exhibition (OFC)*. IEEE, 2017.
- [138] S. Wahls. Generation of time-limited signals in the nonlinear Fourier domain via b-modulation. In *2017 European Conference on Optical Communication (ECOC)*, 2017.
- [139] S. Wahls, S. Chimmalgi, and P. J. Prins. FNFT: A software library for computing nonlinear Fourier transforms. *Journal of Open Source Software*, 3(23):597–598, 2018.
- [140] S. Wahls and H. V. Poor. Introducing the fast nonlinear Fourier transform. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 5780–5784, 2013.
- [141] S. Wahls and H. V. Poor. Fast numerical nonlinear Fourier transforms. *IEEE Transactions on Information Theory*, 61(12):6957–6974, 2015.
- [142] S. Wahls and V. Vaibhav. Fast inverse nonlinear Fourier transforms for continuous spectra of Zakharov-Shabat type. *arXiv preprint arXiv:1607.01305*, 2016.
- [143] S. Watanabe and M. Shirasaki. Exact compensation for both chromatic dispersion and Kerr effect in a transmission fiber using optical phase conjugation. *Journal of Lightwave Technology*, 14(3):243–248, 1996.

- [144] J. Yang. *Nonlinear Waves in Integrable and Nonintegrable systems*. SIAM, 2010.
- [145] X. Yangzhang, V. Aref, S. T. Le, H. Bülow, and P. Bayvel. 400 Gbps dual-polarisation nonlinear frequency-division multiplexed transmission with b-modulation. In *Proc. 2018 European Conference on Optical Communication (ECOC)*, 2018.
- [146] X. Yangzhang, V. Aref, S. T. Le, H. Bülow, D. Lavery, and P. Bayvel. Dual-polarization non-linear frequency-division multiplexed transmission with b-modulation. *J. Lightwave Technol.*, 37(6):1570–1578, 2019.
- [147] X. Yangzhang, S. T. Le, V. Aref, H. Bülow, D. Lavery, and P. Bayvel. Experimental demonstration of dual-polarization NFDm transmission with b-modulation. *IEEE Photon. Technol. Lett.*, 31(11):885–888, 2019.
- [148] M. P. Yankov, F. Da Ros, E. P. da Silva, T. Fehenberger, L. Barletta, D. Zibar, L. K. Oxenløwe, M. Galili, and S. Forchhammer. Nonlinear phase noise compensation in experimental WDM systems with 256QAM. *Journal of Lightwave Technology*, 35(8):1438–1443, 2017.
- [149] A. Yariv, D. Fekete, and D. M. Pepper. Compensation for channel dispersion by nonlinear optical phase conjugation. *Optics Letters*, 4(2):52–54, 1979.
- [150] L. Yi, T. Liao, L. Huang, L. Xue, P. Li, and W. Hu. Machine learning for 100 Gb/s/ $\lambda$  passive optical network. *Journal of Lightwave Technology*, 37(6):1621–1630, 2019.
- [151] M. I. Yousefi and F. R. Kschischang. Information transmission using the Non-linear Fourier Transform, Parts I-III. *IEEE Transactions on Information Theory*, 60(7):4346–4369, 2014.
- [152] H. C. Yuen and B. M. Lake. Nonlinear dynamics of deep-water gravity waves. *Adv. Appl. Mech.*, 22:67–229, 1982.
- [153] N. J. Zabusky and M. D. Kruskal. Interaction of "solitons" in a collisionless plasma and the recurrence of initial states. *Physical Review Letters*, 15(6):240, 1965.
- [154] V. E. Zakharov and A. B. Shabat. Exact theory of 2-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. *Sov. Phys.-JETP*, 34:62–69, 1972.
- [155] D. Zibar, M. Piels, R. Jones, and C. G. Schäffer. Machine learning techniques in optical communication. *Journal of Lightwave Technology*, 34(6):1442–1452, 2015.
- [156] D. Zibar and C. G. Schäffer. Machine learning concepts in coherent optical communication systems. In *Signal Processing in Photonic Communications*, page ST2D1. Optical Society of America, 2014.

# Appendix A

## Fourier transform in MATLAB

MATLAB functions `fft` and `ifft` computes Fourier transform in the manner, when  $Y=\text{fft}(X)$  and  $X=\text{ifft}(Y)$ , both are arrays of length  $N$ :

$$Y_k = \sum_{j=0}^{N-1} X_j e^{-2i\pi jk/N}, \quad X_j = \frac{1}{N} \sum_{k=0}^{N-1} Y_k e^{2i\pi jk/N}, \quad j, k = 0..N-1. \quad (\text{A.1})$$

It is purely invertible, so  $\text{ifft}(\text{fft}(X))=X$ . The first observation here is that this transform is a sum, so, to make it being a discrete kind of integration, we need to multiply it with resolution. The second observation is that for this kind of transform it is natural to have a sampling of variables, starting from zero. Therefore, this transforms can be used directly for the sampling of continuous Fourier transform as follows:

$$\begin{aligned} Y(y) &= \text{FT}(X(x)) = \int_0^{L_x} X(x) e^{-ixy} dx \mapsto \\ \mapsto Y(y_k) &= \sum_j \Delta x e^{-ix_j y_k} X(x_j) = \Delta x \cdot \text{fft}(X(x_j)), \end{aligned} \quad (\text{A.2})$$

where  $x$  and  $y$  are sampled as  $x_k = k\Delta x$  and  $y_j = j \cdot 2\pi/N/\Delta x$ . For inverse transform

$$\begin{aligned} X(x) &= \text{IFT}(Y(y)) = \int_0^{L_y} Y(y) e^{ixy} dy \mapsto \\ \mapsto X(x_k) &= \sum_j \Delta y e^{ix_j y_k} Y(y_k) = \Delta y N \cdot \text{ifft}(Y(y_k)). \end{aligned} \quad (\text{A.3})$$

Something strange happens here with normalisation on the extends  $L_x$  and  $L_y$ , but at least this normalisation is easy to be recovered up to the particular problem. The more complicated story happens when we have another sampling of variables.

In particular, when above we calculated transform from  $l(\xi)$  to  $\Omega_L(\hat{\tau})$ , it was implied that  $\xi$  is sampled symmetrically, when  $\hat{\tau}$  is not. In particular, when we sample the variable  $\xi$  as  $\xi_k = -\Xi/2 + k\Xi/(N-1)$ , so, it has extend  $\Xi$  and resolution  $\Xi/(N-1)$ . Let us write a general relation, when  $x_j = -L_x/2 + jL_x/(N-1)$  and  $y$  is sampled from  $y = 0$ , so,

$y_k = k\Delta y$ , where  $\Delta y$  is unknown yet:

$$\begin{aligned}
 Y(y) &= \int_{-L_x/2}^{L_x/2} X(x)e^{-xy} dx \mapsto \\
 \mapsto Y(y_k) &= \sum_j \Delta x X(x_j) e^{-ix_j y_k} = \\
 &= \sum_j \Delta x X(x_j) \exp\left(-i\left(-\frac{L_x}{2} + j\frac{L_x}{N-1}\right)k\Delta y\right) = \\
 &= \sum_j \Delta x X(x_j) \exp\left(-ij\frac{L_x}{N-1}k\Delta y\right) \exp\left(i\frac{L_x}{2}k\Delta y\right) =
 \end{aligned} \tag{A.4}$$

In order to satisfy relation in exponent in `MATLAB` command, we need  $L_x\Delta y/(N-1) = 2\pi/N$ , so  $\Delta y = 2\pi(N-1)/L_x/N$ . Then the last exponent, as it is sampled over  $k$ , can be moved out of the sum sign and, substituting relation for  $\Delta y$ , one can get:

$$= \frac{L_x}{N-1} \exp\left(\pi k i \frac{N-1}{N}\right) \mathbf{fft}(X(x_j)), \quad k, j = 0..N-1. \tag{A.5}$$

Two last factors both are arrays of length  $N$  and must be multiplied element-wise (`.*` in `MATLAB`).

The rather more natural story, if one has symmetrically sampled both variables. Let say,  $x_j = -L_x/2 + jL_x/(N-1)$  and, as we already got,  $y_k = -\pi(N-1)^2/N/L_x + 2\pi k(N-1)/N/L_x$ , as the step sizes must be related. Subsequently, one gets:

$$\begin{aligned}
 Y(y) &= \int_{-L_x/2}^{L_x/2} dx X(x) e^{-ixy} \mapsto \\
 \mapsto Y(y_k) &= \sum_j \Delta x X(x_j) e^{-ix_j y_k} = \\
 &= \sum_j \Delta x X(x_j) \exp\left(-i\left(-\frac{L_x}{2} + j\frac{L_x}{N-1}\right)\left(-\pi\frac{(N-1)^2}{NL_x} + 2\pi k\frac{N-1}{NL_x}\right)\right) = \\
 &= \sum_j \Delta x X(x_j) \exp\left(-i\frac{L_x}{2}\pi\frac{(N-1)^2}{NL_x}\right) \exp\left(i\frac{L_x}{2}2\pi k\frac{N-1}{NL_x}\right) \times \\
 &\times \exp\left(ij\frac{L_x}{N-1}\pi\frac{(N-1)^2}{NL_x}\right) \exp\left(-ij\frac{L_x}{N-1}2\pi k\frac{N-1}{NL_x}\right) =
 \end{aligned} \tag{A.6}$$

The last exponent appears as a part of `fft`. The first and second exponents can be moved out of sum, but the third one cannot because it is sampled over  $j$ . So, we put it as an input of `fft`:

$$\begin{aligned}
 &= \Delta x \exp\left(-i\frac{\pi}{2}\frac{(N-1)^2}{N}\right) \exp\left(i\pi k\frac{N-1}{N}\right) \sum_j X(x_j) \exp\left(ij\pi\frac{N-1}{N}\right) \exp(-ijk2\pi/N) = \\
 &= \Delta x \exp\left(-i\frac{\pi}{2}\frac{N-1}{N}\right) \exp\left(i\pi k\frac{N-1}{N}\right) \mathbf{fft}\left(X(x_j) \cdot \exp\left(ij\pi\frac{N-1}{N}\right)\right).
 \end{aligned} \tag{A.7}$$

# Appendix B

## Topical visuals

I have dedicated a great deal of my time and efforts during a PhD study to creating bright and informative figures, schemes and animations for topic-related publications and presentations. Despite the ideologically attractive idea of NFT-based optical communications, I strongly believe that it is not communicated in a sufficiently convincing way to the engineering community. I cannot claim, perhaps, that those figures would revolutionary change the situation, though I have received appreciated feedback from colleagues and collaborators. For future use and convenient referencing, here I include a tidy structured collection of those figures.

I combine similar figures with slight variations, as they can highlight different aspects of the presented ideas.

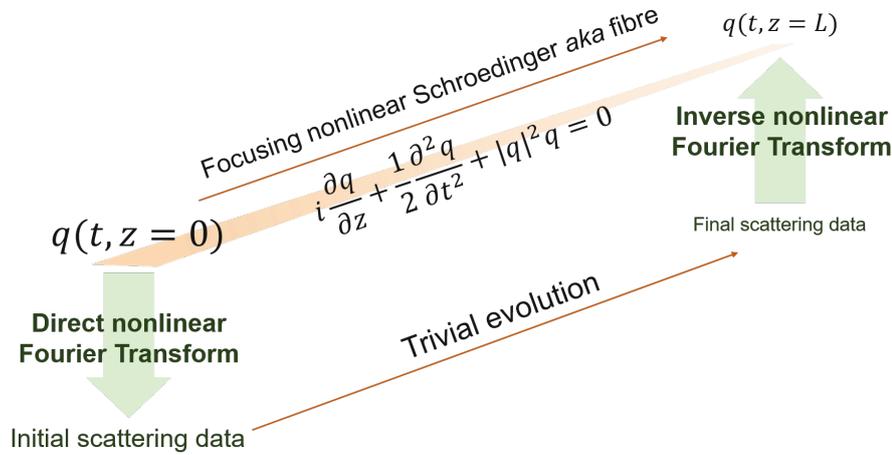


Figure B.1: NLS solution via NFT - I

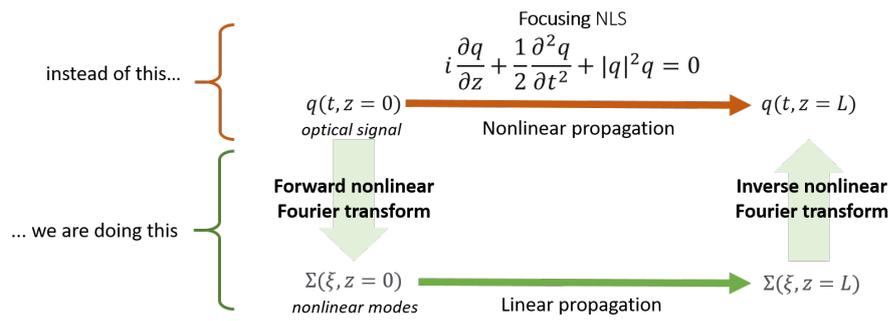


Figure B.2: NLS solution via NFT - II

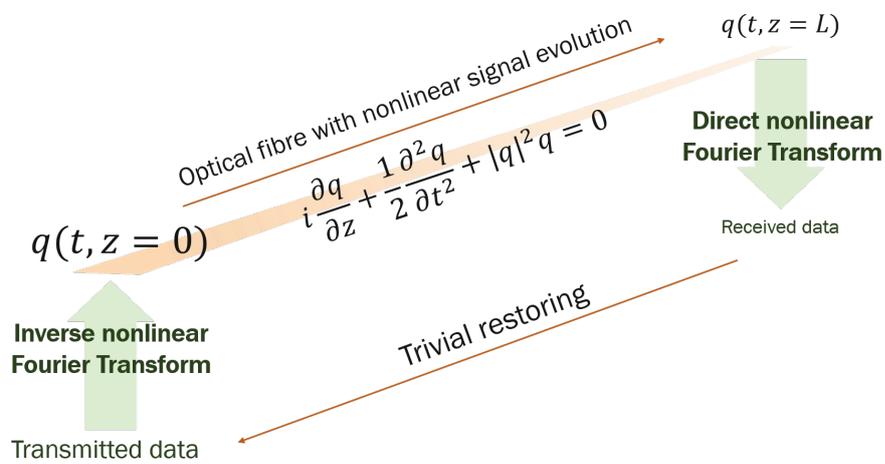


Figure B.3: NFT-based transmission - I

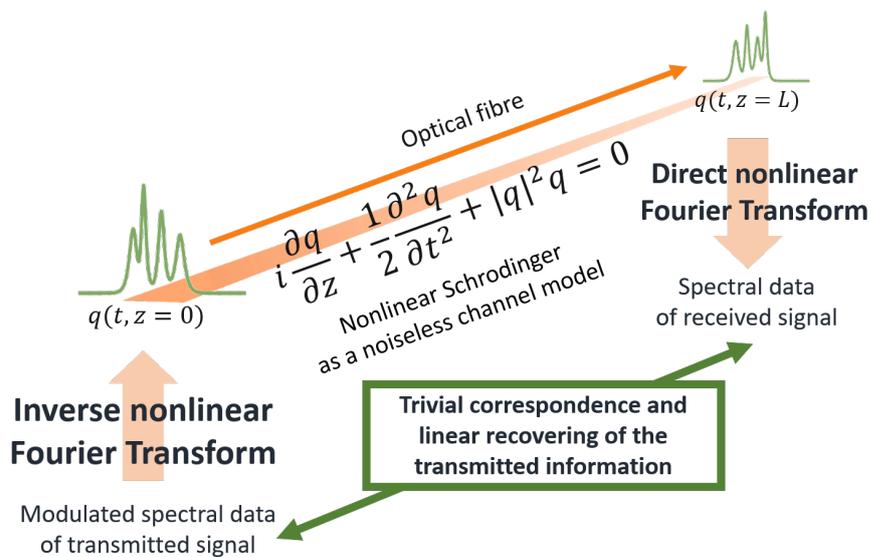


Figure B.4: NFT-based transmission - II

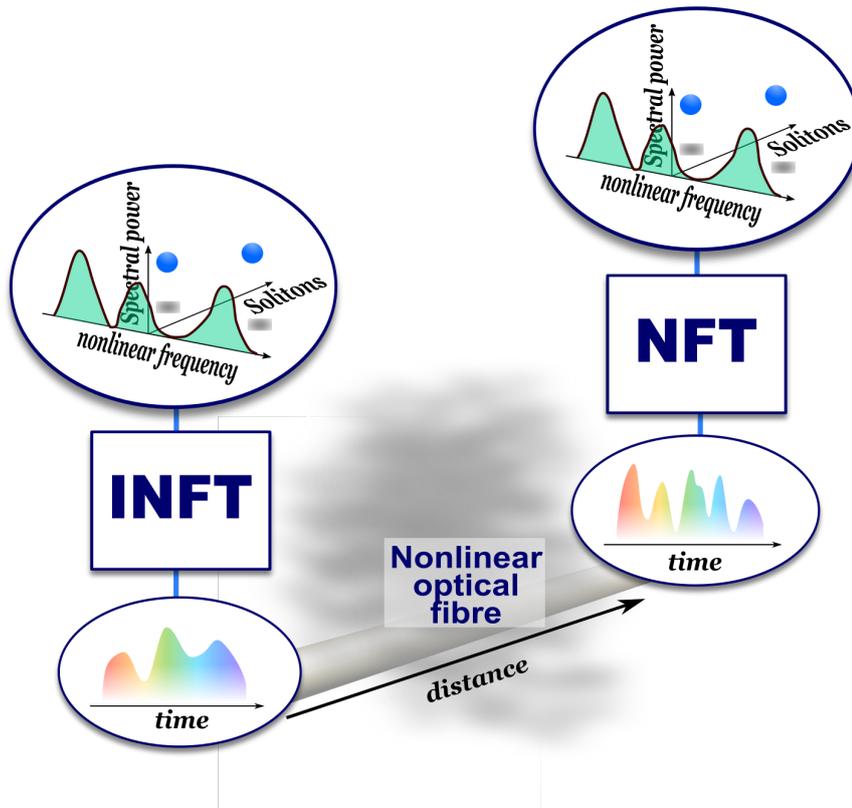


Figure B.5: NFT-based transmission - III

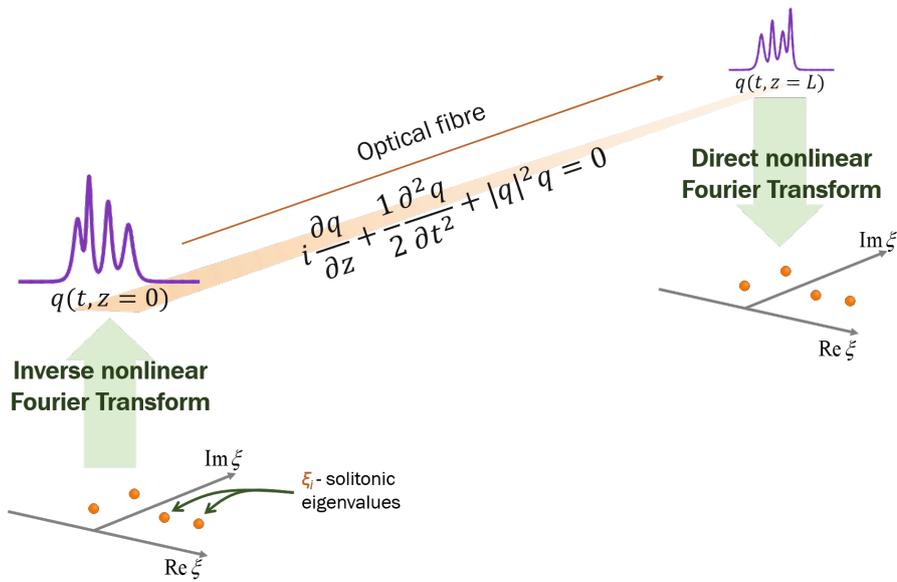


Figure B.6: NFT-based soliton transmission

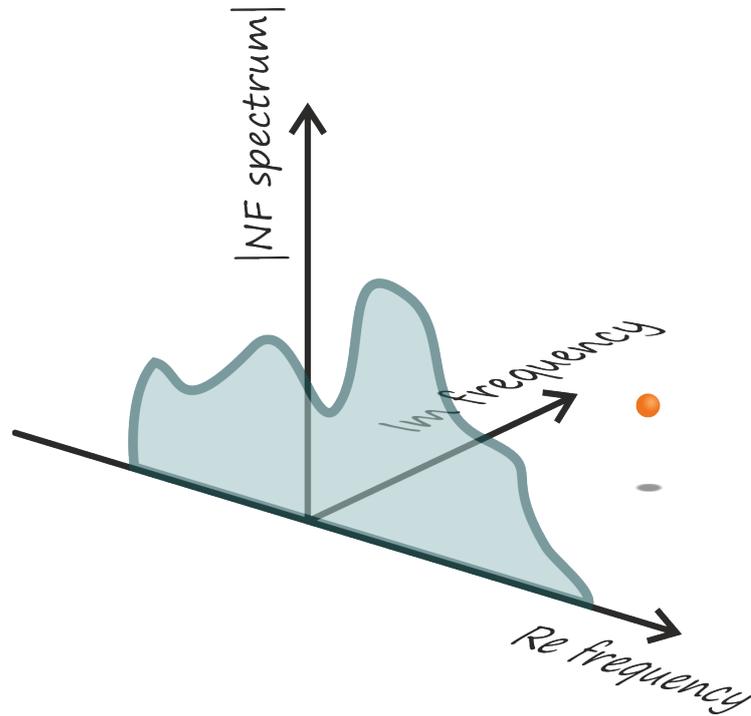


Figure B.7: NF spectrum for vanishing boundary conditions

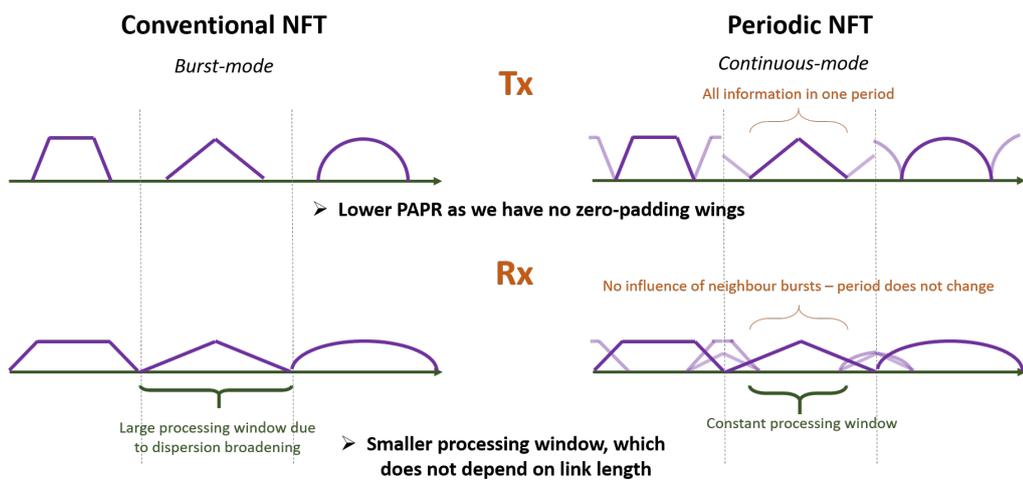


Figure B.8: PNFT vs (vanishing) NFT on Tx and Rx - I

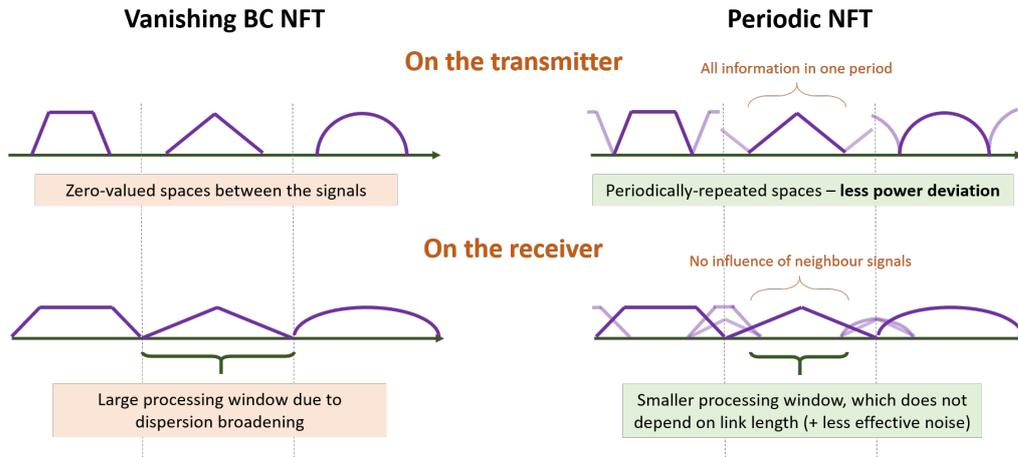


Figure B.9: PNFT vs (vanishing) NFT on Tx and Rx - II

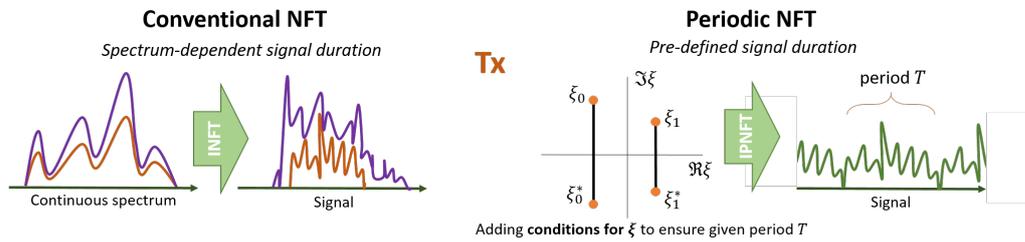


Figure B.10: PNFT vs (vanishing) NFT on Tx - I

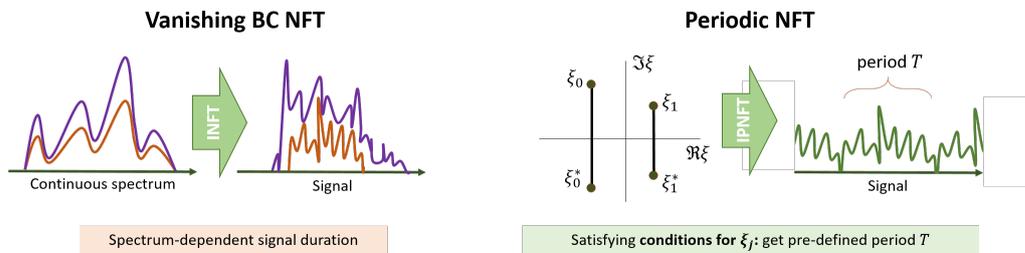


Figure B.11: PNFT vs (vanishing) NFT on Tx - II

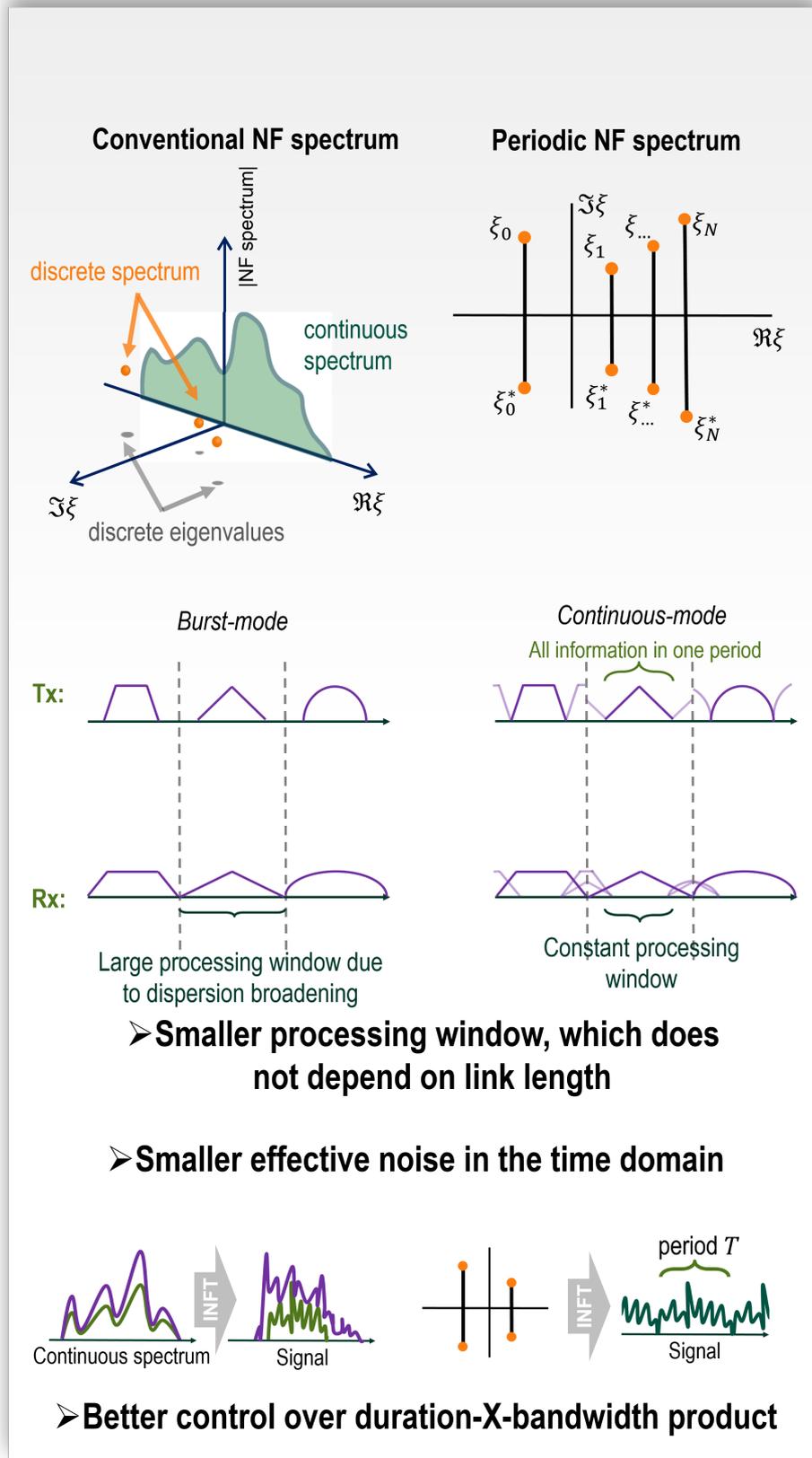


Figure B.12: PNFT vs (vanishing) NFT full comparison

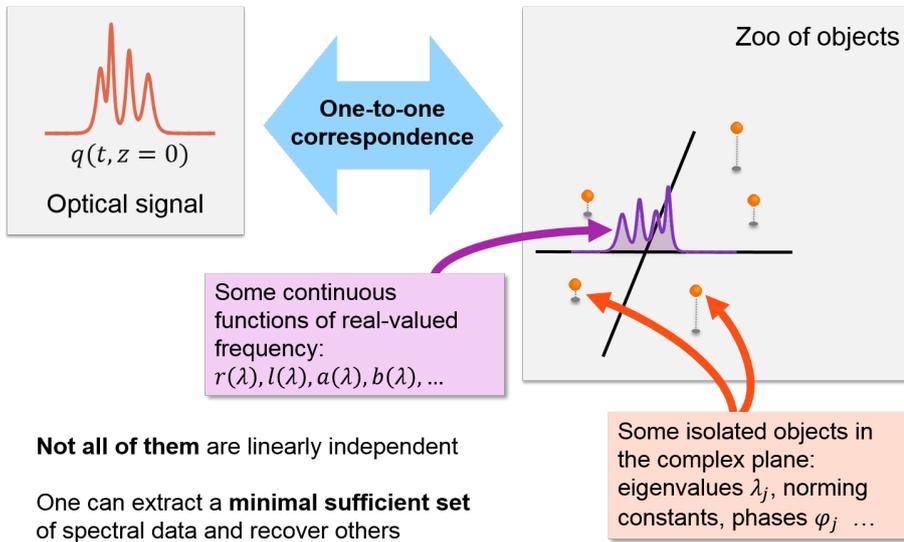


Figure B.13: Details of NF domain - time domain mapping

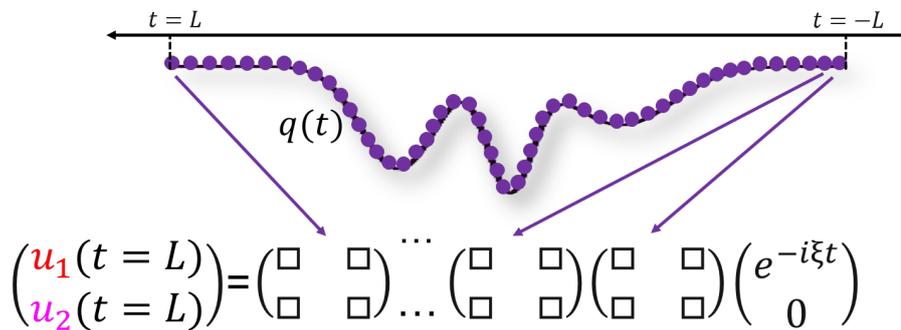


Figure B.14: Transfer matrices approach to ZSP solution